

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2004, 11:19:50 ; Search time 3675 Seconds
(without alignments)
9918.773 Million cell updates/sec

Title: US-09-890-463-5
Perfect score: 841
Sequence: 1 tcggtatcgttaaacagat.....aaaagcgccgctgaatta 841

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.on.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.nam.*
- 37: em.htg.vrt.*
- 38: em.sv.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	840.6	100.0	841	6	BD248905	BD248905 Pigment P
2	824.6	98.0	841	6	BD248906	BD248906 Pigment P
3	702.4	83.5	881	3	AF383156	AF383156 Gonlopora
4	651.4	77.5	693	6	AX699793	AX699793 Sequence
5	650.4	77.3	660	6	AX699783	AX699783 Sequence
6	648.8	77.1	660	6	AX699753	AX699753 Sequence
7	648.8	77.1	660	6	AX699763	AX699763 Sequence
8	648.8	77.1	660	6	AX699785	AX699785 Sequence
9	648.8	77.1	669	6	AX699923	AX699923 Sequence
10	647.6	77.0	663	6	AX699755	AX699755 Sequence
11	647.2	77.0	660	6	AX699759	AX699759 Sequence
12	640.8	76.2	660	6	AX699781	AX699781 Sequence
13	637.6	75.8	660	6	AX699811	AX699811 Sequence
14	637.6	75.8	669	6	AX699921	AX699921 Sequence
15	636	75.6	669	6	AX699925	AX699925 Sequence
16	636	75.6	669	6	AX699927	AX699927 Sequence
17	636	75.6	669	6	AX699933	AX699933 Sequence
18	634.4	75.4	669	6	AX699929	AX699929 Sequence
19	633	75.3	663	6	AX699871	AX699871 Sequence
20	633	75.3	663	6	AX699873	AX699873 Sequence
21	631.4	75.1	663	6	AX699827	AX699827 Sequence
22	631.4	75.1	663	6	AX699829	AX699829 Sequence
23	631.4	75.1	765	6	AX699819	AX699819 Sequence
24	631.4	75.1	765	6	AX699821	AX699821 Sequence
25	628.4	74.7	660	6	AX699823	AX699823 Sequence
26	626.6	74.5	660	6	AX699825	AX699825 Sequence
27	626.4	74.5	660	6	AX699771	AX699771 Sequence
28	626.4	74.5	660	6	AX699773	AX699773 Sequence
29	626.4	74.5	660	6	AX699777	AX699777 Sequence
30	626.4	74.5	660	6	AX699803	AX699803 Sequence
31	626.4	74.5	660	6	AX699815	AX699815 Sequence
32	626.4	74.5	660	6	AX699817	AX699817 Sequence
33	625	74.3	663	6	AX699835	AX699835 Sequence
34	625	74.3	663	6	AX699847	AX699847 Sequence
35	625	74.3	663	6	AX699893	AX699893 Sequence
36	624.8	74.3	660	6	AX699779	AX699779 Sequence
37	623.4	74.1	660	6	AX699899	AX699899 Sequence
38	623.4	74.1	663	6	AX699831	AX699831 Sequence
39	623.4	74.1	663	6	AX699833	AX699833 Sequence
40	623.4	74.1	663	6	AX699879	AX699879 Sequence
41	623.2	74.1	660	6	AX699775	AX699775 Sequence
42	623.2	74.1	660	6	AX699805	AX699805 Sequence
43	623.2	74.1	660	6	AX699809	AX699809 Sequence
44	623.2	74.1	660	6	AX699813	AX699813 Sequence
45	623	74.1	660	6	AX699769	AX699769 Sequence

ALIGNMENTS

RESULT 1
BD248905
LOCUS BD248905 841 bp DNA linear PAT 17-JUL-2003
DEFINITION Pigment protein from coral tissue.
ACCESSION BD248905.1 GI:33058675
VERSION BD248905.1 GI:33058675
KEYWORDS JP 2002535978-A/1.
SOURCE Acropora aspera
ORGANISM Acropora aspera
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
Astrocoenina; Acroporidae; Acropora.
REFERENCE 1 (bases 1 to 841)
AUTHORS Guldborg, O.H. and Dove, S.
TITLE Pigment protein from coral tissue
JOURNAL Patent: JP 2002535978-A 1 29-OCT-2002;

```
COMMENT
OS THE UNIVERSITY OF SYDNEY
PN JP 2002535978-A/1
PF 29-OCT-2000
PR 02-FEB-2000 JP 2000597303
PI OVE HOEGH GULDBERG, SOPHIE DOVE
PC C12N15/09, A61K14/435, C09B61/00, C09K11/06, C12N1/15, PC
C12N1/19,
PC
C12N1/21, C12N5/10, C12P21/02, C12Q1/68, G01N21/78, C12N15/00, C12N5/ PC
00
CC Pigment protein from coral tissue
FH Key Location/Qualifiers
FT source 1..841
FEATURES
source Location/Qualifiers
1..841
/organism="Acropora aspera (plate coral)".
ORIGIN
Query Match 100.0%; Score 840.6; DB 6; Length 841;
Best Local Similarity 100.0%; Pred. No. 5.1e-209;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGACGACCGTCAATGGACAC 60
DB 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGACGACCGTCAATGGACAC 60
QY 61 TACTTTGAGGTGCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGACGGTAAGG 120
DB 61 TACTTTGAGGTGCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGACGGTAAGG 120
QY 121 CTGCTGTGTCACCAAGGGGACCTCTGCATTTGCTTGGATATTTATACACACAGTGT 180
DB 121 CTGCTGTGTCACCAAGGGGACCTCTGCATTTGCTTGGATATTTATACACACAGTGT 180
QY 181 CAGTACGGAAGCATACCATTCACCAAGTACCTGAAGACATCCCTGACTATGTAAGCAG 240
DB 181 CAGTACGGAAGCATACCATTCACCAAGTACCTGAAGACATCCCTGACTATGTAAGCAG 240
QY 241 TCATTCCCGGGAGATATACATGAGGAGAGATCATGAATTTTGAAGATGTCAGTGTGT 300
DB 241 TCATTCCCGGGAGATATACATGAGGAGAGATCATGAATTTTGAAGATGTCAGTGTGT 300
QY 301 ACTGTCAGCAATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAGTTCTCT 360
DB 301 ACTGTCAGCAATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAGTTCTCT 360
QY 361 GGTTTGAACCTTCCTCCAAATGGACCTGTTATGAGAAAGACACACAGGCTGGGAACCC 420
DB 361 GGTTTGAACCTTCCTCCAAATGGACCTGTTATGAGAAAGACACACAGGCTGGGAACCC 420
QY 421 AACCTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGAAAACAACTTTATGCTCTG 480
DB 421 AACCTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGAAAACAACTTTATGCTCTG 480
QY 481 AAGTTAGAAGAGTGTGCTATTTGTCGAATTCAAATCTACTTACAGGCAAGGAG 540
DB 481 AAGTTAGAAGAGTGTGCTATTTGTCGAATTCAAATCTACTTACAGGCAAGGAG 540
QY 541 CCTGTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGATGTAACCAATCAAC 600
DB 541 CCTGTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGATGTAACCAATCAAC 600
QY 601 AAGATTACACTTCGTTGAGCAGCGTGAATTTCCATTGACGCAAACTTTGGTGC 660
DB 601 AAGATTACACTTCGTTGAGCAGCGTGAATTTCCATTGACGCAAACTTTGGTGC 660
QY 661 TGCTGTTTTTTCAGAGTCAAAATCAAGGCACAAATTAAGCAGTGGCTGTAAGAACGTAGATT 720
DB 661 TGCTGTTTTTTCAGAGTCAAAATCAAGGCACAAATTAAGCAGTGGCTGTAAGAACGTAGATT 720
```

```
DB 661 TGCTGTTTTTTCAGAGTCAAAATCAAGGCACAAATTAAGCAGTGGCTGTAAGAACGTAGATT 720
QY 721 CTGATTTTAGCTTAGAGAGTAGGACGAAAGTGTAGACAACTTCAATGATTAAACT 780
DB 721 CTGATTTTAGCTTAGAGAGTAGGACGAAAGTGTAGACAACTTCAATGATTAAACT 780
QY 781 TTTGAAACAAACSCCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 840
DB 781 TTTGAAACAAACSCCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 840
QY 841 A 841
DB 841 A 841
RESULT 2
LOCUS BD248906 841 bp DNA linear PAT 17-JUL-2003
DEFINITION Pigment protein from coral tissue.
ACCESSION BD248906
VERSION BD248906.1 GI:33058676
KEYWORDS JP 2002535978-A/2.
SOURCE Acropora aspera
ORGANISM Acropora aspera
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
Astrocoeniina; Acroporidae; Acropora.
REFERENCE
1 (bases 1 to 841)
Guldborg, O.H. and Dove, S.
Pigment protein from coral tissue
Patent: JP 2002535978-A 2 29-OCT-2002;
THE UNIVERSITY OF SYDNEY
OS Acropora aspera (plate coral)
PN JP 2002535978-A/2
PD 29-OCT-2002
PF 02-FEB-2000 JP 2000597303
PR 02-FEB-1999 AU PP 8463
PI OVE HOEGH GULDBERG, SOPHIE DOVE
PC C12N15/09, A61K14/435, C09B61/00, C09K11/06, C12N1/15, PC
C12N1/19,
PC
C12N1/21, C12N5/10, C12P21/02, C12Q1/68, G01N21/78, C12N15/00, C12N5/ PC
00
CC Pigment protein from coral tissue
FH Key Location/Qualifiers
FT source 1..841
/organism="Acropora aspera (plate coral)".
FEATURES
source Location/Qualifiers
1..841
/organism="Acropora aspera"
/mol_type="genomic DNA"
/db_xref="taxon:140238"
ORIGIN
Query Match 98.0%; Score 824.6; DB 6; Length 841;
Best Local Similarity 98.7%; Pred. No. 7.8e-205;
Matches 830; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
QY 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGACGACCGTCAATGGACAC 60
DB 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGACGACCGTCAATGGACAC 60
QY 61 TACTTTGAGGTGCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGACGGTAAGG 120
DB 61 TACTTTGAGGTGCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGACGGTAAGG 120
QY 121 CTGCTGTCTCACCAGGGGACCTCTGCCATTTCTTGGGATATTTATCACCACAGTGT 180
DB 121 CTGCTGTCTCACCAGGGGACCTCTGCCATTTCTTGGGATATTTATCACCACAGTGT 180
QY 181 CAGTACGGAAGCATACCATTCACCAAGTACCTGAAGACATCCCTGACTATGTAAGCAG 240
DB 181 CAGTACGGAAGCATACCATTCACCAAGTACCTGAAGACATCCCTGACTATGTAAGCAG 240
```

241	QY	TCATTTCCCGGGAGATATACATGGGAGAGGATCATGAACCTTTGAAGATGGTCAGTGTGT	300
241	Db	TCATTTCCCGGGAGATATACATGGGAGAGGATCATGAACCTTTGAAGATGGTCAGTGTGT	300
301	QY	ACTGTCAGCAANTGATTCCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCCTCT	360
301	Db	ACTGTCAGCAANTGATTCCAGCATCCAAGGCAACTGTTTCATCTACCATGTCAAGTTCCTCT	360
361	QY	GGTTTGAACCTTTCTCCCAATGGACCTGTTATGCGAAGAAGACACAGGGCTGGGAACCC	420
361	Db	GGTTTGAACCTTTCTCCCAATGGACCTGTTATGCGAAGAAGACACAGGGCTGGGAACCC	420
421	QY	AACACTGAGCGTCTCTTTGCGACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG	480
421	Db	AACACTGAGCGTCTCTTTGCGACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG	480
481	QY	AAAGTTAGAGGAGGTGGTCACTATTTGTGTAATTCATCTTACAGGCAAGGAAG	540
481	Db	AAAGTTAGAGGAGGTGGTCACTATTTGTGTAATTCATCTTACAGGCAAGGAAG	540
541	QY	CCCTGTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATCTAACCAATCAACAAC	600
541	Db	CCCTGTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATCTAACCAATCAACAAC	600
601	QY	AAGGATTACACTTCCCGTTGAGCAGTGTGAAATTTCCATTTGCACGCAAACTGTGTGTGCGC	660
601	Db	AAGGATTACACTTCCCGTTGAGCAGTGTGAAATTTCCATTTGCACGCAAACTGTGTGTGCGC	660
661	QY	TGCTGTTTTTTCAGAGTCAAAATCAAGGCACAAATAGCAGTGGCGTAAATAAAGCTAGATT	720
661	Db	TGCTGTTTTTTCAGAGTCAAAATCAAGGCACAAATAGCAGTGGCGTAAATAAAGCTAGATT	720
721	QY	CTGATTTTAGCTTTAGAGAAGTGTAGCAACGAAGAAGTGTAGCAACCTTCAATGATTAAACT	780
721	Db	CTGATTTTAGCTTTAGAGAAGTGTAGCAACGAAGAAGTGTAGCAACCTTCAATGATTAAACT	780
781	QY	TTTGAATAACCAACSCCAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGAATT	840
781	Db	TTTGAATAACCAACSCCAATAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGAATT	840
841	QY	A 841	
841	Db	A 841	
RESULT 3			
AF383156	LOCUS	881 bp mRNA linear INV 05-NOV-2001	
DEFINITION	Goniopora tenuidens GFP-like chromoprotein mRNA, complete cds.		
ACCESSION	AF383156		
VERSION	AF383156.1	GI:16660127	
KEYWORDS			
SOURCE	Goniopora tenuidens		
ORGANISM	Goniopora tenuidens		
REFERENCE	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia; Fungia; Poritidae; Goniopora.		
AUTHORS	1 (bases 1 to 881) Gurskaya,N.G., Pradkov,A.F., Terskikh,A., Matz,M.V., Labas,Y.A., Martynov,V.I., Yanushevich,Y.G., Lukyanov,K.A. and Lukyanov,S.A. GFP-like chromoproteins as a source of far-red fluorescent proteins FEBS Lett. 507 (1), 16-20 (2001)		
TITLE			
JOURNAL			
MEDLINE	21538626		
PUBMED	11682051		
REFERENCE	2 (bases 1 to 881) Gurskaya,N.G., Lukyanov,K.A., Labas,Y.A. and Lukyanov,S.A. Direct Submission Submitted (21-MAY-2001) Institute of Bioorganic Chemistry, Miklukho-Maklaya 16/10, Moscow 117997, Russia		
JOURNAL			
FEATURES	Location/Qualifiers		
source	1..881 /organism="Goniopora tenuidens" /mol_type="mRNA" /db_xref="taxon:75301"		

```
RESULT 4
AX699793
LOCUS
DEFINITION
Sequence 61 from Patent WO02070703.
ACCESSION
AX699793
VERSION
AX699793.1 GI:29500268
KEYWORDS
Porites murrayensis
SOURCE
Porites murrayensis
ORGANISM
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
Fungulina; Poritidae; Porites.
REFERENCE
1
AUTHORS
Karan, M., Brugliera, F., Mason, J., Jones, E.L., Dove, S.G.,
Hoegh-Guldberg, I.O. and Prescott, M.
TITLE
Cell visual characteristic-modifying sequences
JOURNAL
Patent: WO 02070703-A 61 12-SEP-2002;
NUFARM AUSTRALIA LIMITED (AU) ; The University of Queensland (AU)
FEATURES
Location/Qualifiers
source
1..693
/organism="Porites murrayensis"
/mol_type="unassigned DNA"
/db_xref="taxon:227229"
<1..663
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAD87465.1"
/db_xref="GI:29500269"
/db_xref="REMTREMBL:CAD87465"
/translation="SVIAKOMTYKVMGTVNGHYVEVGGDKGPKYEGEOTVLTVT
KGGPLPFAWDILSPQSOYSIPFYKPEDIPDYVQSPPEGYTWERIMNFPDGAQCTV
SNDISIQNCFTIYNKFSGLNFPNGVMQKKTQGWENTRLERLARDGLIGNFMAL
KLEGGHHYLCFEKSTYKAKPKVPMGPHYHVDRLDVTNNHNDYTSVEQCEISIAKPKV
VA"
ORIGIN
Query Match 77.5%; Score 651.4; DB 6; Length 693;
Best Local Similarity 96.2%; Pred. No. 1.7e-159;
Matches 667; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
Qy 1 TCCGTTATCGTAAACAGATGACCTACAAAGTTTATATGTGAGGCAAGCTCAATGGACAC 60
Db 1 TCCGTTATCGTAAACAGATGACCTACAAAGTTTATATGTGAGGCAAGCTCAATGGACAC 60
Qy 61 TACTTTGAGTTCGAGGCGATGGAAGAAAGAAAGCCTTACGAGGGGAGCAGACGGTAAGG 120
Db 61 TACTTTGAGTTCGAGGCGATGGAAGAAAGAAAGCCTTACGAGGGGAGCAGACGGTAAGG 120
Qy 121 CTGGCTGTCCACCAAGGGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTGT 180
Db 121 CTGACTGTCCACCAAGGGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTGT 180
Qy 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAGCAG 240
Db 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAGCAG 240
Qy 241 TCATTCCCGGGAGATATACATGAGGAGGATCATGAATTTGAAGATGGTGCAGTGTGT 300
Db 241 TCATTCCCGGGAGATATACATGAGGAGGATCATGAATTTGAAGATGGTGCAGTGTGT 300
Qy 301 ACTGTACGATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAGTTCCT 360
Db 301 ACTGTACGATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAGTTCCT 360
Qy 361 GGTGTGAACCTTCTCCCAATGGACCTGTTATGAGAAAGAGACACAGGGCTGGAAACC 420
Db 361 GGTGTGAACCTTCTCCCAATGGACCTGTTATGAGAAAGAGACACAGGGCTGGAAACC 420
Qy 421 AACACTGAGCGTCTCTTTGACAGAGATGGAATGCTGATAGAAACAACCTTATGGCTGT 480
Db 421 AACACTGAGCGTCTCTTTGACAGAGATGGAATGCTGATAGAAACAACCTTATGGCTGT 480
Qy 481 AAGTTAGGAAGGAGTGGTCACTATTGTGTGAATTCAAATCTACTACAGGCAAGGAG 540
```

```
Db 481 AAGTTGGAAGGAGTGGTCACTATTGTGTGAATTCAAATCTACTTACAGGCAAGGAG 540
Qy 541 CCGTGTGAAGATCCAGGGTATCATTATGTTGACCGCAACTGATGTAACCAATCAAC 600
Db 541 CCGTGTGAAGATCCAGGGTATCATTATGTTGACCGCAACTGATGTAACCAATCAAC 600
Qy 601 AAGGATTACACTTCCGTTGAGCAGCGTGAATTTTCATTGACGCAAAACCTTTGGTCGCC 660
Db 601 AAGGATTACACTTCCGTTGAGCAGCGTGAATTTTCATTGACGCAAAACCTTTGGTCGCC 660
Qy 661 TGCTGTTTTTTCAGAGTCAAAATCAAGGCACAAA 693
Db 661 TGACGTTTTTTCAGAGTCAAAATCAAGGCACAAA 693
RESULT 5
AX699783
LOCUS
DEFINITION
Sequence 51 from Patent WO02070703.
ACCESSION
AX699783
VERSION
AX699783.1 GI:29500258
KEYWORDS
unidentified
SOURCE
unidentified
ORGANISM
unclassified.
REFERENCE
1
AUTHORS
Karan, M., Brugliera, F., Mason, J., Jones, E.L., Dove, S.G.,
Hoegh-Guldberg, I.O. and Prescott, M.
TITLE
Cell visual characteristic-modifying sequences
JOURNAL
Patent: WO 02070703-A 51 12-SEP-2002;
NUFARM AUSTRALIA LIMITED (AU) ; The University of Queensland (AU)
FEATURES
Location/Qualifiers
source
1..660
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/feature="unclassified"
<1..660
/feature="unnamed protein product"
/codon_start=1
/protein_id="CAD87460.1"
/db_xref="GI:29500259"
/db_xref="REMTREMBL:CAD87460"
/translation="SVIAKOMTYKVMGTVNGHYVEVGGDKGPKYEGEOTVLTVT
KGGPLPFAWDILSPQSOYSIPFYKPEDIPDYVQSPPEGYTWERIMNFPDGAQCTV
SNDISIQNCFTIYNKFSGLNFPNGVMQKKTQGWENTRLERLARDGLIGNFMAL
KLEGGHHYLCFEKSTYKAKPKVPMGPHYHVDRLDVTNNHNDYTSVEQCEISIAKPKV
VA"
ORIGIN
Query Match 77.3%; Score 650.4; DB 6; Length 660;
Best Local Similarity 99.1%; Pred. No. 3e-159;
Matches 654; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 TCCGTTATCGTAAACAGATGACCTACAAAGTTTATATGTGAGGCAAGCTCAATGGACAC 60
Db 1 TCCGTTATCGTAAACAGATGACCTACAAAGTTTATATGTGAGGCAAGCTCAATGGACAC 60
Qy 61 TACTTTGAGTTCGAGGCGATGGAAGAAAGAAAGCCTTACGAGGGGAGCAGACGGTAAGG 120
Db 61 TACTTTGAGTTCGAGGCGATGGAAGAAAGAAAGCCTTACGAGGGGAGCAGACGGTAAGG 120
Qy 121 CTGGCTGTCCACCAAGGGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTGT 180
Db 121 CTGACTGTCCACCAAGGGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTGT 180
Qy 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAGCAG 240
Db 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAGCAG 240
Qy 241 TCATTCCCGGGAGATATACATGAGGAGGATCATGAATTTGAAGATGGTGCAGTGTGT 300
```

```
Db      241 TCATTCCCGAGGAGTATACATGGAGGAGGATCATGACATTTTGAAGTGTGACGTGTG 300
Qy      301 ACTGTGAGCAATGATTCAGCATCAAGGCAATCTGTTTCATCTACCATGTCAAGTTCCT 360
Db      301 ACTGTGAGCAATGATTCAGCATCAAGGCAATCTGTTTCATCTACCATGTCAAGTTCCT 360
Qy      361 GGTTTGAACCTTCCCTCCCAATGGACCTGTTATGAGAGAGAGACACAGGGCTGGGAACCC 420
Db      361 GGTTTGAACCTTCCCTCCCAATGGACCTGTTATGAGAGAGAGACACAGGGCTGGGAACCC 420
Qy      421 AACACTGAGCGTCTCTTTGCGACGAGATGGAATGCTGATAGGAAACAACTTTATGSCCTG 480
Db      421 AACACTGAGCGTCTCTTTGCGACGAGATGGAATGCTGATAGGAAACAACTTTATGSCCTG 480
Qy      481 AAGTTAGAGAGGTGTCATCTATTTGTGTAATTCAAATCTACTTACAGGCAAGGAAG 540
Db      481 AAGTTAGAGAGGTGTCATCTATTTGTGTAATTCAAATCTACTTACAGGCAAGGAAG 540
Qy      541 CCTGTGAAGATGCCAGGGTATCCTATCTGTGACCGCAAACTGGATTAACCAATCACAAAC 600
Db      541 CCTGTGAAGATGCCAGGGTATCCTATCTGTGACCGCAAACTGGATTAACCAATCACAAAC 600
Qy      601 AAGGATTACACTTCGGTTGAGCAGCGTGAAATTTCCATTGACGCAAACTTTGTCGCC 660
Db      601 AAGGATTACACTTCGGTTGAGCAGCGTGAAATTTCCATTGACGCAAACTTTGTCGCC 660

RESULT 6
AX699753
LOCUS      AX699753
DEFINITION Sequence 21 from Patent WO02070703.
ACCESSION AX699753
VERSION    AX699753.1 GI:29500228
KEYWORDS   Acropora aspera
SOURCE     Acropora aspera
ORGANISM   Acropora aspera
            Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
            Asterozoenina; Acroporidae; Acropora.

REFERENCE
AUTHORS    Karan,M., Brugliera,F., Mason,J., Jones,E.L., Dove,S.G.,
            Hoegh-Guldberg,I.O. and Prescott,M.
TITLE      Cell visual characteristic-modifying sequences
JOURNAL    Patent: WO 02070703-A 21 12-SEP-2002;
            NUFARM AUSTRALIA LIMITED (AU); The University of Queensland (AU)
FEATURES   source
            1..660
            /organism="Acropora aspera"
            /mol_type="unassigned DNA"
            /db_xref="taxon:140238"
            <1..>660
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD87445.1"
            /db_xref="GI:29500228"
            /db_xref="REMBL:CAD87445"
            /translation="SVIAKQMTYKVMSTGVNHYFEVGDGKPKYEGEOTVRLTVT
            KGGPFPFADWILSPQSGSIPTFKYKPEDIPYVKQSPFEGYTWIRIMNFEGACTV
            SNDSIQNCFLYHVKFSGLNPPGPMQKKTQGWENPTEBLSARDGMLGNFMAL
            KLEGGGHVLCFEKSTYKARKPKVMPFGYHYVDRLDVTNHNKDYTSVEQREISIAKPKV
            VA"

CDS
Query Match 77.1%; Score 648.8; DB 6; Length 660;
Best Local Similarity 98.9%; Pred. No. 7.9e-159;
Matches 653; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy      1 TCGGTTATCGCTAAACAGATGACCTACAAGTTTATATGTGCGGACGGTCAATGACAC 60
Db      1 TCGGTTATCGCTAAACAGATGACCTACAAGTTTATATGTGCGGACGGTCAATGACAC 60
Qy      61 TACTTTGAGTCCGAGCGCATGGAAGAAAGCAACCTTACGAGGGGAGCAGACGGTAAGG 120
Db      61 TACTTTGAGTCCGAGCGCATGGAAGAAAGCAACCTTACGAGGGGAGCAGACGGTAAGG 120
```

```
Qy      121 CTGGCTGTCCCRAGGGGGACCTCTGCCATTTCTTGGGATATTTTATCACCACAGTGT 180
Db      121 CTGACTGTCAACAGGGGGACCTCTGCCATTTCTTGGGATATTTTATCACCACAGTCA 180
Qy      181 CAGTACGGAAGCATACCATTCACCAAGTACCTCTGAAGACATCCCTCTGACTATGTAAGCAG 240
Db      181 CAGTACGGAAGCATACCATTCACCAAGTACCTCTGAAGACATCCCTCTGACTATGTAAGCAG 240
Qy      241 TCATTCCCGGAGGAGATATACATGGGAGAGGATCATGAATTTGAAGATGGTGCAGTGT 300
Db      241 TCATTCCCGGAGGAGATATACATGGGAGAGGATCATGAATTTGAAGATGGTGCAGTGT 300
Qy      301 ACTGTGACGAATGATTCAGCATCAAGGCAATCTGTTTCATCTACCATGTCAAGTTCCT 360
Db      301 ACTGTGACGAATGATTCAGCATCAAGGCAATCTGTTTCATCTACCATGTCAAGTTCCT 360
Qy      361 GGTTTGAACCTTCCCTCCCAATGGACCTGTTATGAGAGAGAGACACAGGGCTGGGAACCC 420
Db      361 GGTTTGAACCTTCCCTCCCAATGGACCTGTTATGAGAGAGAGACACAGGGCTGGGAACCC 420
Qy      421 AACACTGAGCGTCTCTTTGCGACGAGATGGAATGCTGATAGGAAACAACTTTATGSCCTG 480
Db      421 AACACTGAGCGTCTCTTTGCGACGAGATGGAATGCTGATAGGAAACAACTTTATGSCCTG 480
Qy      481 AAGTTAGAGAGGTGTCATCTATTTGTGTAATTCAAATCTACTTACAGGCAAGGAAG 540
Db      481 AAGTTAGAGAGGTGTCATCTATTTGTGTAATTCAAATCTACTTACAGGCAAGGAAG 540
Qy      541 CCTGTGAAGATGCCAGGGTATCCTATCTGTGACCGCAAACTGGATTAACCAATCACAAAC 600
Db      541 CCTGTGAAGATGCCAGGGTATCCTATCTGTGACCGCAAACTGGATTAACCAATCACAAAC 600
Qy      601 AAGGATTACACTTCGGTTGAGCAGCGTGAAATTTCCATTGACGCAAACTTTGTCGCC 660
Db      601 AAGGATTACACTTCGGTTGAGCAGCGTGAAATTTCCATTGACGCAAACTTTGTCGCC 660

RESULT 7
AX699763
LOCUS      AX699763
DEFINITION Sequence 31 from Patent WO02070703.
ACCESSION AX699763
VERSION    AX699763.1 GI:29500238
KEYWORDS   Caulastrea furcata
SOURCE     Caulastrea furcata
ORGANISM   Caulastrea furcata
            Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
            Faviina; Faviidae; Caulastreae.

REFERENCE
AUTHORS    Karan,M., Brugliera,F., Mason,J., Jones,E.L., Dove,S.G.,
            Hoegh-Guldberg,I.O. and Prescott,M.
TITLE      Cell visual characteristic-modifying sequences
JOURNAL    Patent: WO 02070703-A 31 12-SEP-2002;
            NUFARM AUSTRALIA LIMITED (AU); The University of Queensland (AU)
FEATURES   source
            1..660
            /organism="Caulastrea furcata"
            /mol_type="unassigned DNA"
            /db_xref="taxon:46696"
            <1..>660
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAD87450.1"
            /db_xref="GI:29500239"
            /db_xref="REMBL:CAD87450"
            /translation="SVIAKQMTYKVMSTGVNHYFEVGDGKPKYEGEOTVRLTVT
            KGGPFPFADWILSPQSGSIPTFKYKPEDIPYVKQSPFEGYTWIRIMNFEGACTV
            SNDSIQNCFLYHVKFSGLNPPGPMQKKTQGWENPTEBLSARDGMLGNFMAL
            KLEGGGHVLCFEKSTYKARKPKVMPFGYHYVDRLDVTNHNKDYTSVEQREISIAKPKV
            VA"

CDS
ORIGIN
```

```
Query Match      77.1%; Score 648.8; DB 6; Length 660;
Best Local Similarity 98.9%; Pred. No. 7.9e-159;
Matches 653; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGACCGTCAATGGACAC 60
Db 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGACCGTCAATGGACAC 60
QY 61 TACTTTGAGGTGCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGACGGTAAAGG 120
Db 61 TACTTTGAGGTGCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGACGGTAAAGG 120
QY 121 CTGGCTGTCAACCAAGGGGAGACCTCTGCCATTTGCTTGGGATATTTTATCACCAAGTGT 180
Db 121 CTGACTGTCAACCAAGGGGAGACCTCTGCCATTTGCTTGGGATATTTTATCACCAAGTGT 180
QY 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGAGACATCCCTGAGTATTAAGCAG 240
Db 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGAGACATCCCTGAGTATTAAGCAG 240
QY 241 TCATTCCCGGGAGATATACATGCGAGAGGATCATGAACCTTTGAAGATGGTGCAGTGTGT 300
Db 241 TCATTCCCGGGAGATATACATGCGAGAGGATCATGAACCTTTGAAGATGGTGCAGTGTGT 300
QY 301 ACTGTCAGCAATGATTCAGCATCCAGGCAACTGTTTCACTACATCACTGTCAGTGTCTCT 360
Db 301 ACTGTCAGCAATGATTCAGCATCCAGGCAACTGTTTCACTACATCACTGTCAGTGTCTCT 360
QY 361 GGTTTGAACCTTCCCTCCCAATGGACCTGTTATGCAAGAAAGACACAGGGCTGGGAACCC 420
Db 361 GGTTTGAACCTTCCCTCCCAATGGACCTGTTATGCAAGAAAGACACAGGGCTGGGAACCC 420
QY 421 AACACTGAGCGTCTCTTTCGACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG 480
Db 421 AACACTGAGCGTCTCTTTCGACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG 480
QY 481 AAGTTAGAGAGGCTGTCTCTTTCGACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG 540
Db 481 AAGTTAGAGAGGCTGTCTCTTTCGACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG 540
QY 541 CCTGTGAAGATGCCAGGGTATACATGTTGACCGCAAACTGGATGAACCAATCACAAAC 600
Db 541 CCTGTGAAGATGCCAGGGTATACATGTTGACCGCAAACTGGATGAACCAATCACAAAC 600
QY 601 AAGGATTACACTTCGCTTGGAGCGGTGAAATTTCCATTGACGCAAACTTTGTCGCC 660
Db 601 AAGGATTACACTTCGCTTGGAGCGGTGAAATTTCCATTGACGCAAACTTTGTCGCC 660

RESULT 8
AX699785
LOCUS      AX699785
DEFINITION Sequence 53 from Patent W002070703.
ACCESSION AX699785
VERSION    AX699785.1 GI:29500260
KEYWORDS   unidentified
SOURCE     unidentified
ORGANISM   unclassified.

REFERENCE 1
AUTHORS    Karan,M., Brugliera,F., Mason,J., Jones,E.L., Dove,S.G.,
            Hoegh-Guldberg,I.O. and Prescott,M.
TITLE      Cell visual characteristic-modifying sequences
JOURNAL    Patent: WO 02070703-A 53 12-SEP-2002;
            NUFARM AUSTRALIA LIMITED (AU) ; The University of Queensland (AU)
FEATURES   Location/Qualifiers
            1..660
            /organism="unidentified"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32644"
            /note="Millepora sp. (Hydrozoan)"
            <1..>660
            /note="unnamed protein product"

CDS
            1..660
            /note="unnamed protein product"

ORIGIN
Query Match      77.1%; Score 648.8; DB 6; Length 660;
Best Local Similarity 98.9%; Pred. No. 7.9e-159;
Matches 653; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGACCGTCAATGGACAC 60
Db 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGACCGTCAATGGACAC 60
QY 61 TACTTTGAGGTGCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGACGGTAAAGG 120
Db 61 TACTTTGAGGTGCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGACGGTAAAGG 120
QY 121 CTGGCTGTCAACCAAGGGGAGACCTCTGCCATTTGCTTGGGATATTTTATCACCAAGTGT 180
Db 121 CTGACTGTCAACCAAGGGGAGACCTCTGCCATTTGCTTGGGATATTTTATCACCAAGTGT 180
QY 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGAGACATCCCTGAGTATTAAGCAG 240
Db 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGAGACATCCCTGAGTATTAAGCAG 240
QY 241 TCATTCCCGGGAGATATACATGCGAGAGGATCATGAACCTTTGAAGATGGTGCAGTGTGT 300
Db 241 TCATTCCCGGGAGATATACATGCGAGAGGATCATGAACCTTTGAAGATGGTGCAGTGTGT 300
QY 301 ACTGTCAGCAATGATTCAGCATCCAGGCAACTGTTTCACTACATCACTGTCAGTGTCTCT 360
Db 301 ACTGTCAGCAATGATTCAGCATCCAGGCAACTGTTTCACTACATCACTGTCAGTGTCTCT 360
QY 361 GGTTTGAACCTTCCCTCCCAATGGACCTGTTATGCAAGAAAGACACAGGGCTGGGAACCC 420
Db 361 GGTTTGAACCTTCCCTCCCAATGGACCTGTTATGCAAGAAAGACACAGGGCTGGGAACCC 420
QY 421 AACACTGAGCGTCTCTTTCGACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG 480
Db 421 AACACTGAGCGTCTCTTTCGACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG 480
QY 481 AAGTTAGAGAGGCTGTCTCTTTCGACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG 540
Db 481 AAGTTAGAGAGGCTGTCTCTTTCGACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG 540
QY 541 CCTGTGAAGATGCCAGGGTATACATGTTGACCGCAAACTGGATGAACCAATCACAAAC 600
Db 541 CCTGTGAAGATGCCAGGGTATACATGTTGACCGCAAACTGGATGAACCAATCACAAAC 600
QY 601 AAGGATTACACTTCGCTTGGAGCGGTGAAATTTCCATTGACGCAAACTTTGTCGCC 660
Db 601 AAGGATTACACTTCGCTTGGAGCGGTGAAATTTCCATTGACGCAAACTTTGTCGCC 660

RESULT 9
AX699923
LOCUS      AX699923
DEFINITION Sequence 191 from Patent W002070703.
ACCESSION AX699923
VERSION    AX699923.1 GI:29500381
KEYWORDS   Discosoma sp.
SOURCE     Discosoma sp.
ORGANISM   Discosoma sp.
            Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia;
            Corallimorpharia; Discosomatidae; Discosoma.
REFERENCE 1
AUTHORS    Karan,M., Brugliera,F., Mason,J., Jones,E.L., Dove,S.G.,
```

```

Hoeigh-Guldberg, I.O. and Prescott, M.
Cell visual characteristic-modifying sequences
Patent: WO 02070703-A 191 12-SEP-2002;
NUFARM AUSTRALIA LIMITED (AU) ; The University of Queensland (AU)
FEATURES
    source
        1..669
        /organism="Discosoma sp."
        /mol_type="unassigned DNA"
        /db_xref="taxon:86600"
ORIGIN
    Query Match      77.1%; Score 648.8; DB 6; Length 669;
    Best Local Similarity 98.9%; Pred. No. 7.9e-159;
    Matches 653; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 1 TCCGTTATCGCTAAACAGATGACCTTACAAAGTTTATATGTCAGGCAACGGTCAATGGACAC 60
Db 4 TCCGTTATCGCTAAACAGATGACCTTACAAAGTTTATATGTCAGGCAACGGTCAATGGACAC 63
Qy 61 TACTTTGAGGTCGAGGCGATGGAAGAGAGCCCTTACGAGGGGAGCAGCAGGTAAGG 120
Db 64 TACTTTGAGGTCGAGGCGATGGAAGAGAGCCCTTACGAGGGGAGCAGCAGGTAAGG 123
Qy 121 CTGGCTGTCAACAGGGCGGACCTCTGCCAATTTGCTTGGGATATTTATCACCACAGTGT 180
Db 124 CTGACTGTCAACAGGGCGGACCTCTGCCAATTTGCTTGGGATATTTATCACCACAGTCA 183
Qy 181 CAGTACGGAAGCATACCATTCACCAAGTACCTTGAAGACATCCCTGACTATGTAAGCAG 240
Db 184 CAGTACGGAAGCATACCATTCACCAAGTACCTTGAAGACATCCCTGACTATGTAAGCAG 243
Qy 241 TCATTTCCCGGAGGATATACATGCGGAGGATCATGAATTTTGAAGATGGTCAAGTGTCT 300
Db 244 TCATTTCCCGGAGGATATACATGCGGAGGATCATGAATTTTGAAGATGGTCAAGTGTCT 303
Qy 301 ACTGTCAGCAATGATTCAGCATCCAAAGCAACTGTTTCATCTACCATGTCGAAGTTCTCT 360
Db 304 ACTGTCAGCAATGATTCAGCATCCAAAGCAACTGTTTCATCTACCATGTCGAAGTTCTCT 363
Qy 361 GGTTTGAATTTCTCCCAATGCACTGTTATGCGAAGAACACACAGGGCTGGGAACCC 420
Db 364 GGTTTGAATTTCTCCCAATGCACTGTTATGCGAAGAACACACAGGGCTGGGAACCC 423
Qy 421 AACACTGAGCGTCTTTTCACGAGATGGAATGCTCATAGGAACAACTTTATGGCTCTG 480
Db 424 AACACTGAGCGTCTTTTCACGAGATGGAATGCTCATAGGAACAACTTTATGGCTCTG 483
Qy 481 AAGTTAGAGGAGGTGCTCACTATTGTTGTAATTCAAATCTACTTACAGGCAAGGAAG 540
Db 484 AAGTTAGAGGAGGTGCTCACTATTGTTGTAATTCAAATCTACTTACAGGCAAGGAAG 543
Qy 541 CTTGTAAGATGCGAGGGTATCACTATGTTGACCGCAAACTGGATGTAAACCAATCAAC 600
Db 544 CTTGTAAGATGCGAGGGTATCACTATGTTGACCGCAAACTGGATGTAAACCAATCAAC 603
Qy 601 AAGGATTACACTCCGTCAGCAGCGTGAATTTCCATTGCGAGCAAACTTTGGTCGCC 660
Db 604 AAGGATTACACTCCGTCAGCAGCGTGAATTTCCATTGCGAGCAAACTTTGGTCGCC 663

RESULT 10
LOCUS AX699755
DEFINITION Sequence 23 from Patent WO02070703.
ACCESSION AX699755
VERSION AX699755.1 GI:29500230
KEYWORDS
SOURCE Acropora aspera
ORGANISM Acropora aspera
            Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Scleractinia;
            Astrocoeniina; Acroporidae; Acropora.
REFERENCE 1
AUTHORS Karan, M., Brugliera, F., Mason, J., Jones, E.L., Dove, S.G.,

```

```

RESULT 11
AX699759
LOCUS AX699759 660 bp DNA linear PAT 02-APR-2003
DEFINITION Sequence 27 from Patent WO02070703.
ACCESSION AX699759
VERSION AX699759.1 GI:29500234
KEYWORDS Acanthastrea echinata
SOURCE Acanthastrea echinata
ORGANISM Acanthastrea echinata
Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Scleractinia;
Favaria; Mussidae; Acanthastrea.
1
Karan,M., Brugliera,F., Mason,J., Jones,E.L., Dove,S.G.,
Hoegh-Guldberg,I.O. and Prescott,M.
TITLE Cell visual characteristic-modifying sequences
JOURNAL Patent: WO 02070703-A 27 12-SEP-2002;
NUFARM AUSTRALIA LIMITED (AU) ; The University of Queensland (AU)
FEATURES
source
1..660
/organism="Acanthastrea echinata"
/mol_type="unassigned DNA"
/db_xref="taxon:51051"
<1..>660
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD87448.1"
/db_xref="GI:29500235"
/db_xref="REMBL:CAD87448"
/translation="SVIAKQMTYKVMSTGVNHYFEVGGKPKYEGEQTVLRTVT
KGGPLPAWDLSPQSYGSIPTFKYEDIPDYVKQSPFEGYTWERINFPDGAQCTV
SNDSIQNCFIYHVKFSLFPNPGVPMQKTOGWEPNRLERLARDQMLIGNFMAL
KLEGGHYLCEFKSTYKARKPKVMPGVHCVDRKLDVTNNHNDYTSVEGREISIAKPV
VA"
CDS
Query Match 77.0%; Score 647.2; DB 6; Length 660;
Best Local Similarity 98.8%; Pred. No. 2.1e-158;
Matches 652; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 TCCGTTATCGCTAACAGATCACTACAAAGTTTATATGTCAGGACCGTCAATGGAC 60
Db 1 TCCGTTATCGCTAACAGATCACTACAAAGTTTATATGTCAGGACCGTCAATGGAC 60
QY 61 TACTTTGAGGTCGAAGGCGATGGAAGAGCCTTACGAGGGGAGCAGACGGTAAGG 120
Db 61 TACTTTGAGGTCGAAGGCGATGGAAGAGCCTTACGAGGGGAGCAGACGGTAAGG 120
QY 121 CTGGCTGTACCAAGGGCGACCTCTGCCATTTGCTGGGATATTTATACCAAGTGT 180
Db 121 CTGACTGTACCAAGGGCGACCTCTGCCATTTGCTGGGATATTTATACCAAGTCA 180
QY 181 CAGTACGGAAGCATACCATTCACCAAGTACCTGAAGACATCCCTGACTATGTAAGCAG 240
Db 181 CAGTACGGAAGCATACCATTCACCAAGTACCTGAAGACATCCCTGACTATGTAAGCAG 240
QY 241 TCATTCCCGGGAGGATATACATGGGAGGATCATGAATTTGAAGATGTCGAGTGTGT 300
Db 241 TCATTCCCGGGAGGATATACATGGGAGGATCATGAATTTGAAGATGTCGAGTGTGT 300
QY 301 ACTGTACGATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAAGTCTCT 360
Db 301 ACTGTACGATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAAGTCTCT 360
QY 361 GGTTTGAACCTTCTCCCAATGGACCTGTTATGTCAGAGAGAGACACAGGGCTGGGAACC 420
Db 361 GGTTTGAACCTTCTCCCAATGGACCTGTTATGTCAGAGAGAGACACAGGGCTGGGAACC 420
QY 421 AACACTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGGAAACAATTTATGGCTCTG 480
Db 421 AACACTGAGCGTCTCTCTGACGAGATGGAATGCTGATAGGAAACAATTTATGGCTCTG 480

```

```

QY 481 AAGTTAGAGGAGGTGGTCACTATTGTTGTAATTCAAATCTACTTTACAAGCAAGGAAG 540
Db 481 AAGTTAGAGGAGGTGGTCACTATTGTTGTAATTCAAATCTACTTTACAAGCAAGGAAG 540
QY 541 CCGTGAAGATGCCAGGGTATCACTATGTTGACCCCAAACTGGATGTAACCAATCACAAC 600
Db 541 CCGTGAAGATGCCAGGGTATCACTATGTTGACCCCAAACTGGATGTAACCAATCACAAC 600
QY 601 AAGGATTACACTTCGTTGAGCAGCGTGAATTTCCATTGCAACCACTTGTGTCGCC 660
Db 601 AAGGATTACACTTCGTTGAGCAGCGTGAATTTCCATTGCAACCACTTGTGTCGCC 660

RESULT 12
AX699781
LOCUS AX699781 660 bp DNA linear PAT 02-APR-2003
DEFINITION Sequence 49 from Patent WO02070703.
ACCESSION AX699781
VERSION AX699781.1 GI:29500256
KEYWORDS
SOURCE unidentified
ORGANISM unclassified.
1
Karan,M., Brugliera,F., Mason,J., Jones,E.L., Dove,S.G.,
Hoegh-Guldberg,I.O. and Prescott,M.
TITLE Cell visual characteristic-modifying sequences
JOURNAL Patent: WO 02070703-A 49 12-SEP-2002;
NUFARM AUSTRALIA LIMITED (AU) ; The University of Queensland (AU)
FEATURES
source
1..660
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="Willepora sp. (Hydrozoan)"
<1..>660
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD87459.1"
/db_xref="GI:29500257"
/db_xref="REMBL:CAD87459"
/translation="SVIAKQMTYKVMSTGVNHYFEVGGKPKYEGEQTVLRTVT
KGGPLPAWDLSPQSYGSIPTFKYEDIPDYVKQSPFEGYTWERINFPDGAQCTV
SNDSIQNCFIYHVKFSLFPNPGVPMQKTOGWEPNRLERLARDQMLIGNFMAL
KLEGGHYLCEFKSTYKARKPKVMPGVHCVDRKLDVTNNHNDYTSVEGREISIAKPV
VA"
CDS
Query Match 76.2%; Score 640.8; DB 6; Length 660;
Best Local Similarity 98.2%; Pred. No. 9.8e-157;
Matches 648; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
QY 1 TCCGTTATCGCTAACAGATCACTACAAAGTTTATATGTCAGGACCGTCAATGGAC 60
Db 1 TCCGTTATCGCTAACAGATCACTACAAAGTTTATATGTCAGGACCGTCAATGGAC 60
QY 61 TACTTTGAGGTCGAAGGCGATGGAAGAGCCTTACGAGGGGAGCAGACGGTAAGG 120
Db 61 TACTTTGAGGTCGAAGGCGATGGAAGAGCCTTACGAGGGGAGCAGACGGTAAGG 120
QY 121 CTGGCTGTACCAAGGGCGACCTCTGCCATTTGCTGGGATATTTATACCAAGTGT 180
Db 121 CTCACTGTACCAAGGGCGACCTCTGCCATTTGCTGGGATATTTATACCAAGTGT 180
QY 181 CAGTACGGAAGCATACCATTCACCAAGTACCTGAAGACATCCCTGACTATGTAAGCAG 240
Db 181 CAGTACGGAAGCATACCATTCACCAAGTACCTGAAGACATCCCTGACTATGTAAGCAG 240
QY 241 TCATTCCCGGGAGGATATACATGGGAGGATCATGAATTTGAAGATGTCGAGTGTGT 300
Db 241 TCATTCCCGGGAGGATATACATGGGAGGATCATGAATTTGAAGATGTCGAGTGTGT 300
QY 301 ACTGTACGATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAAGTCTCT 360

```

```
Db 301 ACTGTGACGATGATCCAGCATCAAGGCAACTGTTTCATCTACCATGTCAAGTCTCT 360
Qy 361 GGTGTGAACCTTCTCCCAATGGACCTTTATGCGAAGAAAGACACACAGGCTGGAAACC 420
Db 361 GGTGTGAACCTTCTCCCAATGGACCTTTATGCGAAGAAAGACACACAGGCTGGAAACC 420
Qy 421 AACACTGAGCGTCTCTTTGCGAGAGATGGAATGCTATAGGAACAACTTTATGGCTCTG 480
Db 421 AACACTGAGCGTCTCTTTGCGAGAGATGGAATGCTATAGGAACAACTTTATGGCTCTG 480
Qy 481 AAGTTAGAGAGAGTGGTCACTATTTGTGTGAATTCAAATCTACTTCAAGGCAAGGAAG 540
Db 481 AAGTTAGAGAGAGTGGTCACTATTTGTGTGAATTCAAATCTACTTCAAGGCAAGGAAG 540
Qy 541 CCTGTGAAGATGCCAGGATACATATGTTGACCGCAAACTGGATGAACCAATCACAAAC 600
Db 541 CCTGTGAAGATGCCAGGATACATATGTTGACCGCAAACTGGATGAACCAATCACAAAC 600
Qy 601 AAGGATTACACTTCGTTGAGCAGCGTGAAATTTCCATTGCAAGCAAACTGGTGGTGGCC 660
Db 601 AAGGATTACACTTCGTTGAGCAGCGTGAAATTTCCATTGCAAGCAAACTGGTGGTGGCC 660

RESULT 13
AX699811
LOCUS AX699811 660 bp DNA linear PAT 03-APR-2003
DEFINITION Sequence 79 from Patent W002070703.
ACCESSION AX699811
VERSION AX699811.1 GI:29500285
KEYWORDS
SOURCE Pavona decussata
ORGANISM Pavona decussata
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
Fungiini; Agariciidae; Pavona.
REFERENCE
1
AUTHORS Karan,M., Brugliera,F., Mason,J., Jones,E.L., Dove,S.G.,
Hoegh-Guldberg,I.O. and Prescott,M.
TITLE Cell visual characteristic-modifying sequences
JOURNAL Patent: WO 02070703-A 79 12-SEP-2002;
NUPARM AUSTRALIA LIMITED (AU) ; The University of Queensland (AU)
FEATURES
source
1..660
/organism="Pavona decussata"
/mol_type="unassigned DNA"
/db_xref="taxon:227230"
<1..>660
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD87473.1"
/db_xref="GI:29500286"
/db_xref="EMBL:CAD87473"
/translation="SVIAKQMTYKVMGTVNGHYFVEGDKGPKYEGEQTVKLTVT
EGGPLPFWADILSPQSQVSPFTKYPEDIPDYKQSPFEGYTWERIMNFEDGAVCTV
SNDSSIOGNCFTYHKVKSGLNFPNGVMOKTKQWEPNTERLSARDGMLIGNPMAL
KLEGGHYLCFEKSYKARKPKVMPGYHVDRKLDVTHNKNKYTSVEQREISIAKPKV
VA"
CDS
ORIGIN
Query Match 75.8%; Score 637.6; DB 6; Length 660;
Best Local Similarity 97.9%; Pred. No. 6.7e-156;
Matches 646; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 1 TCCGTTATCGCTAAACAGATGACCTTCAAAAGTTTATATGTCAGGCACGGTCAATGACAC 60
Db 1 TCCGTTATCGCTAAACAGATGACCTTCAAAAGTTTATATGTCAGGCACGGTCAATGACAC 60
Qy 61 TACTTTGAGGTGAAAGCCATGGAAGAAAGCCCTTACAGGGGAGCAGCGGTAAGG 120
Db 61 TACTTTGAGGTGAAAGCCATGGAAGAAAGCCCTTACAGGGGAGCAGCGGTAAGG 120
Qy 121 CTGGCTGTGTCACCAAGGGCGGACCTCTGCCATTGCTTGGGATATTTTATCACCACAGTGT 180
```

```
Db 121 CTCACTGTCCAGGCGGACCTCTGCCATTTGCTTGGGATATTTCTATCACCACAGAGT 180
Qy 181 CAGTACGGAACATACCATTCACCAAGTACCTCGAGACATCCCTGACTATGTAAGCAG 240
Db 181 CAGTACGGAACGATACCATTCACCAAGTACCTCGAGACATCCCTGACTATGTAAGCAG 240
Qy 241 TCATTCCCGGGAGATATACATGCGAGAGGATCATGAATTTGAAGATGGTGCAGTGT 300
Db 241 TCATTCCCGGGAGATATACATGCGAGAGGATCATGAATTTGAAGATGGTGCAGTGT 300
Qy 301 ACTGTGAGCAATGATTCAGCATCAAGGCAAACTGTTTTCATCTACCATGTCAAGTCTCT 360
Db 301 ACTGTGAGCAATGATTCAGCATCAAGGCAAACTGTTTTCATCTACCATGTCAAGTCTCT 360
Qy 361 GGTGTGAACCTTCTCCCAATGGACCTTTATGCGAAGAAAGACACACAGGCTGGAAACC 420
Db 361 GGTGTGAACCTTCTCCCAATGGACCTTTATGCGAAGAAAGACACACAGGCTGGAAACC 420
Qy 421 AACACTGAGCGTCTCTTTGCGAGAGATGGAATGCTATAGGAACAACTTTATGGCTCTG 480
Db 421 AACACTGAGCGTCTCTTTGCGAGAGATGGAATGCTATAGGAACAACTTTATGGCTCTG 480
Qy 481 AAGTTAGAGAGAGTGGTCACTATTTGTGTGAATTCAAATCTACTTCAAGCAAGGAAG 540
Db 481 AAGTTAGAGAGAGTGGTCACTATTTGTGTGAATTCAAATCTACTTCAAGCAAGGAAG 540
Qy 541 CCTGTGAAGATGCCAGGATACATATGTTGACCGCAAACTGGATGAACCAATCACAAAC 600
Db 541 CCTGTGAAGATGCCAGGATACATATGTTGACCGCAAACTGGATGAACCAATCACAAAC 600
Qy 601 AAGGATTACACTTCGTTGAGCAGCGTGAAATTTCCATTGCAAGCAAACTGGTGGTGGCC 660
Db 601 AAGGATTACACTTCGTTGAGCAGCGTGAAATTTCCATTGCAAGCAAACTGGTGGTGGCC 660

RESULT 14
AX699921
LOCUS AX699921 669 bp DNA linear PAT 02-APR-2003
DEFINITION Sequence 189 from Patent W002070703.
ACCESSION AX699921
VERSION AX699921.1 GI:29500380
KEYWORDS
SOURCE Acropora aspera
ORGANISM Acropora aspera
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
Astrocoenelina; Acroporidae; Acropora.
REFERENCE
1
AUTHORS Karan,M., Brugliera,F., Mason,J., Jones,E.L., Dove,S.G.,
Hoegh-Guldberg,I.O. and Prescott,M.
TITLE Cell visual characteristic-modifying sequences
JOURNAL Patent: WO 02070703-A 189 12-SEP-2002;
NUPARM AUSTRALIA LIMITED (AU) ; The University of Queensland (AU)
FEATURES
source
1..669
/organism="Acropora aspera"
/mol_type="unassigned DNA"
/db_xref="taxon:140238"
ORIGIN
Query Match 75.8%; Score 637.6; DB 6; Length 669;
Best Local Similarity 97.9%; Pred. No. 6.8e-156;
Matches 646; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Qy 1 TCCGTTATCGCTAAACAGATGACCTTCAAAAGTTTATATGTCAGGCACGGTCAATGACAC 60
Db 4 TCCGTTATCGCTAAACAGATGACCTTCAAAAGTTTATATGTCAGGCACGGTCAATGACAC 63
Qy 61 TACTTTGAGGTGAAAGCCATGGAAGAAAGCCCTTACAGGGGAGCAGCGGTAAGG 120
Db 64 TACTTTGAGGTGAAAGCCATGGAAGAAAGCCCTTACAGGGGAGCAGCGGTAAGG 123
Qy 121 CTGGCTGTGTCACCAAGGGCGGACCTCTGCCATTGCTTGGGATATTTTATCACCACAGTGT 180
```

Db 124 CTCACGTGTCCACAGGGTGGACCTCTGCCATTTGCTGGGATATTTTATCACCACAGTCA 183
QY 181 CAGTACGGAAGCATACCACTTCCACCAAGTACCTCGAAGACATCCCTGACTATGTAAAGCAG 240
Db 184 CAGTACGGAAGCATACCACTTCCACCAAGTACCTCGAAGACATCCCGACTATGTAAAGCAG 243
QY 241 TCATTCCCGGGAGATATACATGAGGAGGATCATGAACCTTTGAAGATGGTGCAGTGTGT 300
Db 244 TCATTCCCGGGAGGATATACATGAGGAGGATCATGAACCTTTGAAGATGGTGCAGTGTGT 303
QY 301 ACTGTGACCAATGATTCAGCATCCAGGCACTGTTTCACTACCAATGTCAGTTCTCT 360
Db 304 ACTGTGACCAATGATTCAGCATCCAGGCACTGTTTCACTACCAATGTCAGTTCTCT 363
QY 361 GGTTTGAACTTTCTCTCCCAATGGACCTGTATGCAAGAGAGACACAGGGCTGGGAACCC 420
Db 364 GGTTTGAACTTTCTCTCCCAATGGACCTGTATGCAAGAGAGACACAGGGCTGGGAACCC 423
QY 421 AACACTGAGCGTCTTTTGCAGAGATGGAATGCTGATAGGAAACAACCTTTATGGCTCTG 480
Db 424 AACACTGAGCGTCTTTTGCAGAGATGGAATGCTGATAGGAAACAACCTTTATGGCTCTG 483
QY 481 AAGTTAGAGAGGTGGTCACTATTGTTGTAATTCBAATCTACTTACAGGCAAGGAAG 540
Db 484 AAGTTAGAGAGGTGGTCACTATTGTTGTAATTCBAATCTACTTACAGGCAAGGAAG 543
QY 541 CCTGTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGAACCAATCACAAC 600
Db 544 CCTGTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGAACCAATCACAAC 603
QY 601 AAGGATTACACTTCGGTTGAGCAGCGTGAATTTCCATTGCAACGCAAACTTTTGGTCGCC 660
Db 604 AAGGATTACACTTCGGTTGAGCAGCGTGAATTTCCATTGCAACGCAAACTTTTGGTCGCC 663

RESULT 15
AX699925
LOCUS AX699925 669 bp DNA linear PAT 02-APR-2003
DEFINITION Sequence 193 from Patent WO02070703.
ACCESSION AX699925
VERSION AX699925.1 GI:29500382
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
unclassified.
REFERENCE 1
AUTHORS Karan,M., Bruggiera,F., Mason,J., Jones,E.L., Dove,S.G.,
Hoegh-Guldberg,I.O. and Prescott,M.
TITLE Cell visual characteristic-modifying sequences
JOURNAL Patent: WO 02070703-A 193 12-SEP-2002;
NUFARM AUSTRALIA LIMITED (AU) ; The University of Queensland (AU)
FEATURES
Location/Qualifiers
source 1..669
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="Simularia sp."

ORIGIN

Query Match 75.6%; Score 636; DB 6; Length 669;
Best Local Similarity 97.7%; Pred. No.1.8e-155;
Matches 645; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
QY 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGAGGACCGTCAATGGACAC 60
Db 4 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGAGGACCGTCAATGGACAC 63
QY 61 TACTTTGAGTTCGAGGCGATGGAAGAGGCTTACGAGGGGAGGACACCGGTAAAG 120
Db 64 TACTTTGAGTTCGAGGCGATGGAAGAGGCTTACGAGGGGAGGACACCGGTAAAG 123
QY 121 CTGGCTGTCCACAAAGGGCGGACCTCTGCCATTTGCTGGGATATTTTATCACCACAGTGT 180

Search completed: August 13, 2004, 19:12:32
Job time : 3676 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2004, 11:19:50 ; Search time 409.5 Seconds

(Without alignments)
8724.627 Million cell updates/sec

Title: US-09-890-463-6

Perfect score: 841

Sequence: 1 tccgttatcctaacaagat.....aaaagcgccgcgaatta 841

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 29Jan04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002s:*

7: geneseqn2003as:*

8: geneseqn2003bs:*

9: geneseqn2003cs:*

10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	841	100.0	841	3	AA52083 Pigment P
2	824.6	98.0	841	3	AA52082 Pigment P
3	657.8	78.2	693	6	AB212089 Colour Fa
4	652.4	77.6	663	6	AB212070 Colour Fa
5	648.8	77.1	660	6	AB212084 Colour Fa
6	647.2	77.0	660	6	AB212085 Colour Fa
7	647.2	77.0	660	6	AB212074 Colour Fa
8	647.2	77.0	660	6	AB212069 Colour Fa
9	647.2	77.0	669	6	AB212156 Colour Fa
10	645.6	76.8	660	6	AB212072 Colour Fa
11	642.4	76.4	669	6	AB212155 Colour Fa
12	640.8	76.2	669	6	AB212161 Colour Fa
13	640.8	76.2	669	6	AB212158 Colour Fa
14	639.2	76.0	660	6	AB212083 Colour Fa
15	639.2	76.0	669	6	AB212159 Colour Fa
16	637.8	75.8	663	6	AB212129 Colour Fa
17	637.8	75.8	663	6	AB212128 Colour Fa
18	637.6	75.8	669	6	AB212157 Colour Fa
19	636.2	75.6	663	6	AB212107 Colour Fa
20	636.2	75.6	663	6	AB212106 Colour Fa
21	636.2	75.6	765	6	AB212103 Colour Fa
22	636.2	75.6	765	6	AB212102 Colour Fa
23	636	75.6	660	6	AB212098 Colour Fa

24	633.2	75.3	660	6	AB212104	Colour Fa
25	631.4	75.1	660	6	AB212105	Colour Fa
26	631.2	75.1	660	6	AB212100	Colour Fa
27	631.2	75.1	660	6	AB212079	Colour Fa
28	631.2	75.1	660	6	AB212078	Colour Fa
29	631.2	75.1	660	6	AB212094	Colour Fa
30	631.2	75.1	660	6	AB212081	Colour Fa
31	631.2	75.1	660	6	AB212101	Colour Fa
32	625.8	74.9	663	6	AB212110	Colour Fa
33	629.8	74.9	663	6	AB212139	Colour Fa
34	629.8	74.9	663	6	AB212116	Colour Fa
35	629.6	74.9	660	6	AB212082	Colour Fa
36	628.2	74.7	660	6	AB212142	Colour Fa
37	628.2	74.7	663	6	AB212132	Colour Fa
38	628.2	74.7	663	6	AB212108	Colour Fa
39	628.2	74.7	663	6	AB212109	Colour Fa
40	628	74.7	660	6	AB212080	Colour Fa
41	628	74.7	660	6	AB212097	Colour Fa
42	628	74.7	660	6	AB212095	Colour Fa
43	627.8	74.6	660	6	AB212077	Colour Fa
44	627.4	74.6	660	6	AB212075	Colour Fa
45	626.6	74.5	660	6	AB212135	Colour Fa

ALIGNMENTS

RESULT 1

AA52083
ID AA52083 standard; cDNA; 841 BP.

AC AA52083;

DT 04-DEC-2000 (first entry)

DE Pigment protein from coral tissue POC4 cDNA.

KW N-terminal; pigment protein from coral tissue; PPCT; fluorescence;

KW tissue marker; fluorescent marker; dyestuff; sunscreen; ultra violet;

KW UV filter; POC3; ss.

OS Acropora aspera.

EH Key Location/Qualifiers

FT CDS 1..708

FT /*tag= a

FT /label= POC4

FT /product= "Pigment_protein_from_coral_tissue"

FT /partial

XX WO200046233-A1.

XX 10-AUG-2000.

XX 02-FEB-2000; 2000WO-AU0000056.

XX 02-FEB-1999; 99AU-00008463.

XX (UNSY) UNIV SYDNEY.

XX Hoegh-Guldberg O, Dove S;

XX WPI; 2000-532892/48.

XX P-PSDB; AAY97150.

XX Novel pigment protein derived from corals capable of emitting fluorescence upon irradiation by incident light useful as tissue marker.

XX fluorescent marker or general dyestuff.

XX Claim 10; Page 44; 49pp; English.

XX cDNA libraries were constructed from a blue pigmented coral, Acropora

XX aspera to isolate sequences encoding polypeptides with N-terminal

CC sequences as shown in AAY97147-48. Pigment protein from coral tissue
CC (PPCT) is capable of emitting fluorescence upon irradiation by incident
CC light whose maximal absorbance is in the range of 320-600 nm and a
CC maximal fluorescence emission is in the range of 300-700 nm. PPCT may be
CC used as a tissue marker, fluorescent marker (e.g. to follow gene
CC expression in transformed tissues) or general dyestuff (all claimed).
CC PPCT may also be used in sunscreen formulations or UV filters (both
CC claimed)
XX
SQ Sequence 841 BP; 275 A; 171 C; 195 G; 200 T; 0 U; 0 Other;

Query Match 100.0%; Score 841; DB 3; Length 841;
Best Local Similarity 100.0%; Pred. No. 2.5e-130;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGCAACGTCATCGACAC 60
Db 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGCAACGTCATCGACAC 60

QY 61 TACTTTGAGTTCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGCGTAAAGG 120
Db 61 TACTTTGAGTTCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGCGTAAAGG 120

QY 121 CTGGCTGTACCAAGGGCGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGT 180
Db 121 CTGGCTGTACCAAGGGCGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGT 180

QY 181 CAGTACGGAAGCATACCATTCACCAAGTACCTGAAAGACATCCCTGACTATGTAAGCAG 240
Db 181 CAGTACGGAAGCATACCATTCACCAAGTACCTGAAAGACATCCCTGACTATGTAAGCAG 240

QY 241 TCATTCCCGGGAGATATACATGGGAGGATCATGAACCTTTGAAAGATGGTGCAGTGT 300
Db 241 TCATTCCCGGGAGATATACATGGGAGGATCATGAACCTTTGAAAGATGGTGCAGTGT 300

QY 301 ACTGTCAGCATGATTCAGCATCAAGGCACTGTTTCATCTACCATGTCAGTTCTCT 360
Db 301 ACTGTCAGCATGATTCAGCATCAAGGCACTGTTTCATCTACCATGTCAGTTCTCT 360

QY 361 GGTGTGAACCTTCTCCCAATGGACCTGTTATGCAGAGAGACACAGGGCTGGAAACC 420
Db 361 GGTGTGAACCTTCTCCCAATGGACCTGTTATGCAGAGAGACACAGGGCTGGAAACC 420

QY 421 AACCTAGAGCTCTCTTTGACGAGATGGAATGCTATAGGAAACAACTTTATGGCTCTG 480
Db 421 AACCTAGAGCTCTCTTTGACGAGATGGAATGCTATAGGAAACAACTTTATGGCTCTG 480

QY 481 AAGTTAGAGAGGTGGTCACTATTGTTGTAATTCAAATCTACTTACAGGCAAGAG 540
Db 481 AAGTTAGAGAGGTGGTCACTATTGTTGTAATTCAAATCTACTTACAGGCAAGAG 540

QY 541 CCTGTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAC 600
Db 541 CCTGTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAC 600

QY 601 AAGATTACACTTCGTTGACAGTGTGAATTTCCATTGTCAGCAAAACCTGTGTGCGC 660
Db 601 AAGATTACACTTCGTTGACAGTGTGAATTTCCATTGTCAGCAAAACCTGTGTGCGC 660

QY 661 TGCCGTTTTTTCAGAGTCAAAATCAAGGCACAAATACGAGTGGCGTAAACACGTAGATT 720
Db 661 TGCCGTTTTTTCAGAGTCAAAATCAAGGCACAAATACGAGTGGCGTAAACACGTAGATT 720

QY 721 CTGATTTTATAGCTTATAGAAAGTGAACGAAGTGTAAACACCACTTAATGATTAACT 780
Db 721 CTGATTTTATAGCTTATAGAAAGTGAACGAAGTGTAAACCACTTAATGATTAACT 780

QY 781 TTTGAAACACAGCCATAAAAAATAAAAAAATAAAAAAATAAAAAAAGCGCGCTCGAATT 840
Db 781 TTTGAAACACAGCCATAAAAAATAAAAAAATAAAAAAATAAAAAAAGCGCGCTCGAATT 840

841 A 841

Db 841 A 841

RESULT 2
AA52082
ID AA52082 standard; cDNA; 841 BP.
XX
AC AA52082;
XX
DT 04-DEC-2000 (first entry)
XX
DE Pigment protein from coral tissue POC3 cDNA.
XX
KW N-terminal; pigment protein from coral tissue; PPCT; fluorescence;
KW tissue marker; fluorescent marker; dyestuff; sunscreen; ultra violet;
KW UV filter; POC3; ss.
XX
OS Acropora aspera.
XX
FH Key Location/Qualifiers
CDS 1..696
FT /*tag= a
FT /label= POC3
FT /product= "Pigment_protein_from_coral_tissue"
FT /partial
XX
PN WO200046233-A1.
XX
PD 10-AUG-2000.
XX
PF 02-FEB-2000; 2000WO-AU0000056.
XX
PR 02-FEB-1999; 99AU-00008463.
XX
PA (UNSY) UNIV SYDNEY.
XX
PI Hoegh-Guldberg O, Dove S;
XX
DR WPI; 2000-532892/48.
XX
P-PSDB; AAY97149.
XX
PT Novel pigment protein derived from corals capable of emitting
PT fluorescence upon irradiation by incident light useful as tissue marker,
PT fluorescent marker or general dyestuff.
XX
PS Claim 10; Page 44; 49pp; English.
XX
CC cDNA libraries were constructed from a blue pigmented coral, Acropora
CC aspera to isolate sequences encoding polypeptides with N-terminal
CC sequences as shown in AAY97147-48. Pigment protein from coral tissue
CC (PPCT) is capable of emitting fluorescence upon irradiation by incident
CC light whose maximal absorbance is in the range of 320-600 nm and a
CC maximal fluorescence emission is in the range of 300-700 nm. PPCT may be
CC used as a tissue marker, fluorescent marker (e.g. to follow gene
CC expression in transformed tissues) or general dyestuff (all claimed).
CC PPCT may also be used in sunscreen formulations or UV filters (both
CC claimed)
SQ Sequence 841 BP; 274 A; 171 C; 196 G; 199 T; 0 U; 1 Other;

Query Match 98.0%; Score 824.6; DB 3; Length 841;
Best Local Similarity 98.7%; Pred. No. 2e-186;
Matches 830; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGCAACGTCATCGACAC 60
Db 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGCAACGTCATCGACAC 60

QY 61 TACTTTGAGTTCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGCGTAAAGG 120
Db 61 TACTTTGAGTTCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGCGTAAAGG 120

QY 121 CTGGCTGTACCAAGGGCGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGT 180

```
Db 121 CTGCTGTCACCAAGGGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTGT 180
Qy 181 CAGTACGGAAGCATACCAATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAAGCAG 240
Db 181 CAGTACGGAAGCATACCAATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAAGCAG 240
Qy 241 TCATTCGCGGGAGATATACATGGGAGGAGATCATGAATTTGAGATGTTGACGTGTGT 300
Db 241 TCATTCGCGGGAGATATACATGGGAGGAGATCATGAATTTGAGATGTTGACGTGTGT 300
Qy 301 ACTGTCACATGATTCAGCATCCCAAGCACTGTTTCACTACCATGTCAAGTTCTCT 360
Db 301 ACTGTCACATGATTCAGCATCCCAAGCACTGTTTCACTACCATGTCAAGTTCTCT 360
Qy 361 GGTGTGAATTTCTTCCCAATGGACCTGTTATGCAAGAGAGACACAGGGGTGGAAACC 420
Db 361 GGTGTGAATTTCTTCCCAATGGACCTGTTATGCAAGAGAGACACAGGGGTGGAAACC 420
Qy 421 AACACTGAGCGCTCTTTTGCAGGAGATGGAATGCTGATAGGAACAATTTATGCTCTG 480
Db 421 AACACTGAGCGCTCTTTTGCAGGAGATGGAATGCTGATAGGAACAATTTATGCTCTG 480
Qy 481 AAGTTAGAGGAGGTGCTCACTATTTGTGAATTCAAATCTACTTACAGGCAAGGAAG 540
Db 481 AAGTTAGAGGAGGTGCTCACTATTTGTGAATTCAAATCTACTTACAGGCAAGGAAG 540
Qy 541 CCGTGAAGATGCCAGGATATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAAC 600
Db 541 CCGTGAAGATGCCAGGATATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAAC 600
Qy 601 AAGGATTACACTTCGTTGACGAGTGTGAATTTCCATTTGACGCAAACTGTTGTCGCC 660
Db 601 AAGGATTACACTTCGTTGACGAGTGTGAATTTCCATTTGACGCAAACTGTTGTCGCC 660
Qy 661 TGCCGCTTTTTCAGAGTCAAAATCAAGGACAAATACGAGTGGCTGCTTAAAAACGTAGATT 720
Db 661 TGCTGTTTTTCAGAGTCAAAATCAAGGACAAATACGAGTGGCTGCTTAAAAACGTAGATT 720
Qy 721 CTGATTTTAGCTTATAGAGTAGGACGAGAGAGTGTAAACCACTTAATGATTAAGT 780
Db 721 CTGATTTTAGCTTATAGAGTAGGACGAGAGAGTGTAAACCACTTAATGATTAAGT 780
Qy 781 TTTGAAACACGCCATATATATATATATATATATATATATATATATATATATATATATAT 840
Db 781 TTTGAAACACGCCATATATATATATATATATATATATATATATATATATATATATATAT 840
Qy 841 A 841
Db 841 A 841

RESULT 3
ID ABZ12089 standard; DNA; 693 BP.
XX AC ABZ12089;
XX
XX
XX 22-JAN-2003 (first entry)
XX DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 61.
XX
XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
XX UV sink; sunscreen; ds.
XX
XX Porites murrayensis.
XX
XX WO200270703-A2.
XX
XX 12-SEP-2002.
XX
```

```
PF 01-MAR-2002; 2002WO-GB000928.
XX
XX 02-MAR-2001; 2001US-0273227P.
PR 21-MAR-2001; 2001AU-00003874.
PR 15-OCT-2001; 2001US-0329816P.
XX
XX (NUFA-) NUFARM LTD.
PA (UYOU ) UNIV QUEENSLAND.
PA (JONE/) JONES E L.
PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
PI Hoegh-Guldberg IO, Prescott M;
XX WPI; 2002-740765/80.
XX
XX Novel color-facilitating molecule for producing a biomatrix, has a
XX polypeptide which alone/along with molecules imparts altered visual
XX characteristics to cells in the absence of excitation by extraneous non-
XX white light.
XX
XX Claim 6; Page 338-339; 510pp; English.
XX
XX The invention relates to an isolated colour-facilitating molecule (CFM)
XX comprising a polypeptide which, in a cell, alone or together with one or
XX more other molecules imparts an altered visual characteristic to the cell
XX when visualised by a human eye in the absence of excitation by extraneous
XX non-white light or particle emission. CFMs are useful for producing a
XX transgenic animal which exhibits a novel colour e.g. sheep with blue or
XX red coloured fleece. They are useful for producing coloured plant
XX extracts, e.g. flavouring, beverage or juice or colouring agent. Other
XX uses include transducing or intensifying an image, providing additional
XX light for growing phototropic organisms e.g. algae and/or corals, for
XX coating materials that experience UV damage e.g. plastics and car
XX upholstery. CFMs are useful in the flower industry, in the development of
XX new varieties of flowering plants. Other contemplated uses include,
XX expression markers, general reporter molecules, photon traps, UV sinks or
XX in sunscreens. CFMs modify visible colour in edible and/or ornamental
XX fungal species, and in fruits and vegetables to enhance their
XX marketability. CFMs embedded in a gel matrix improve image quality in
XX situations of distorted light spectra (biomatrix). The first all-protein
XX chromophore to be isolated was Green Fluorescent protein (GFP). The
XX sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
XX sequences
XX
XX Sequence 693 BP; 204 A; 152 C; 167 G; 170 T; 0 U; 0 Other;
XX
XX Query Match 78.2%; Score 657.8; DB 6; Length 693;
XX Best Local Similarity 96.8%; Pred. No. 9.3e-147;
XX Matches 671; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
XX
Qy 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGACGCTCAATGGACAC 60
Db 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGACGCTCAATGGACAC 60
Qy 61 TACTTTGAGTCCGAAAGCGGATGAAAAAGAAAGCCCTTACAGGGGGAGACGGTAAAG 120
Db 61 TACTTTGAGTCCGAAAGCGGATGAAAAAGAAAGCCCTTACAGGGGGAGACGGTAAAG 120
Qy 121 CTGGGTGTCCCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGT 180
Db 121 CTCACTGTCCCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGT 180
Qy 181 CAGTACGGAAGCATACCAATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAAGCAG 240
Db 181 CAGTACGGAAGCATACCAATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAAGCAG 240
Qy 241 TCATTTCCCGGGAGATATACATGGGAGGAGATCATGAATTTGAGATGTTGACGTGTGT 300
Db 241 TCATTTCCCGGGAGATATACATGGGAGGAGATCATGAATTTGAGATGTTGACGTGTGT 300
Qy 301 ACTGTCAGCAATGATTCAGCATCCCAAGCACTGTTTCACTACCATGTCAAGTTCTCT 360
Db 301 ACTGTCAGCAATGATTCAGCATCCCAAGCACTGTTTCACTACCATGTCAAGTTCTCT 360
Qy 301 ACTGTCAGCAATGATTCAGCATCCCAAGCACTGTTTCACTACCATGTCAAGTTCTCT 360
Db 301 ACTGTCAGCAATGATTCAGCATCCCAAGCACTGTTTCACTACCATGTCAAGTTCTCT 360
```

QY 361 GGTGTAAGAGGAGGCTCTTCCCAATGGACCTGTTATGCAAGAGAGACACACAGGCGTGGAAACC 420
 Db |||||
 QY 421 AACACTGAGCGTCTCTTTGACAGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG 480
 Db |||||
 QY 481 AAGTTAGAGGAGGCTGCTCACTATTGTTGTGAATTCAAATCTACTTACAAAGCAAGAG 540
 Db |||||
 QY 541 CCTGTGAGATGCCAGGGTATCATTGTTGTGACCGCAAACTCGATGTAACCAATCACAA 600
 Db |||||
 QY 601 AAGGATTACACTTCCGTTGAGCAGTGTGAATTCATTTGCAAGCAAACTGCTGCTGCC 660
 Db |||||
 QY 661 TGCGTTTTTTCAGAGTCAAAATCAAGGCACAAA 693
 Db |||||
 QY 661 TGAGTTTTTTCAGAGTCAAAATCAAGGCACAAA 693

RESULT 4

ID ABZ12070 standard; DNA; 663 BP.

AC ABZ12070;

DT 22-JAN-2003 (first entry)

DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 23.

KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunscreen; ds.

OS Acropora aspera.

XX WO200270703-A2.

XX 12-SEP-2002.

XX 01-MAR-2002; 2002WO-GB000928.

XX 02-MAR-2001; 2001US-0273227P.

XX 21-MAR-2001; 2001AU-00003874.

XX 15-OCT-2001; 2001US-0329816P.

XX (NUFA-) NUFARM LTD.

PA (UYQU) UNIV QUEENSLAND.

PA (JONE/) JONES E L.

XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;

XX Hoegh-Guldberg IO, Prescott M;

XX WPI; 2002-740765/80.

XX Novel color-facilitating molecule for producing a biomatrix, has a

XX polypeptide which alone/along with molecules imparts altered visual

XX characteristics to cells in the absence of excitation by extraneous non-

XX white light.

XX Claim 6; Page 287-288; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)

XX comprising a polypeptide which, in a cell, alone or together with one or

XX more other molecules imparts an altered visual characteristic to the cell

XX when visualised by a human eye in the absence of excitation by extraneous

CC non-white light or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant
 CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development of
 CC new varieties of flowering plants. Other contemplated uses include:
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
 CC sequences

XX SQ Sequence 663 BP; 193 A; 145 C; 164 G; 161 T; 0 U; 0 Other;

Query Match 77.6%; Score 652.4; DB 6; Length 663;

Best Local Similarity 99.1%; Pred. No. 1.8e-145;

Matches 556; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TCCGTTATCGCTAAACAGATGACCTTACAAAGTTTATATGTCAGGCACGCTCAATGACAC 60
 Db |||||
 QY 1 TCCGTTATCGCTAAACAGATGACCTTACAAAGTTTATATGTCAGGCACGCTCAATGACAC 60
 Db |||||
 QY 61 TACTTTGAGTTCGAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGCGTAAGG 120
 Db |||||
 QY 61 TACTTTGAGTTCGAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGCGTAAGG 120
 Db |||||
 QY 121 CTGCTGTCCACCAAGGGCGACCTCTGCCATTGCTGGGATATTTTATCACCACAGTGT 180
 Db |||||
 QY 121 CTGCTGTCCACCAAGGGCGACCTCTGCCATTGCTGGGATATTTTATCACCACAGTGT 180
 Db |||||
 QY 181 CAGTACGGAAGCATACCAATTCCAAAGTACCTTGAAGCATCCCTGACTATGTAAGCAG 240
 Db |||||
 QY 181 CAGTACGGAAGCATACCAATTCCAAAGTACCTTGAAGCATCCCTGACTATGTAAGCAG 240
 Db |||||
 QY 241 TCATTCGCGGGAGATATACATGGGAGAGGATCATGAATTTGAAGTGGTGCAGTGTGT 300
 Db |||||
 QY 241 TCATTCGCGGGAGATATACATGGGAGAGGATCATGAATTTGAAGTGGTGCAGTGTGT 300
 Db |||||
 QY 301 ACTGTTCAGCATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT 360
 Db |||||
 QY 301 ACTGTTCAGCATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT 360
 Db |||||
 QY 361 GGTGTAAGTCTTCCCAATGGACCTGTTATGAGAAAGAGACACAGGGCTGGAAACC 420
 Db |||||
 QY 361 GGTGTAAGTCTTCCCAATGGACCTGTTATGAGAAAGAGACACAGGGCTGGAAACC 420
 Db |||||
 QY 421 AACACTGAGCGTCTCTTTGACAGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG 480
 Db |||||
 QY 421 AACACTGAGCGTCTCTTTGACAGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG 480
 Db |||||
 QY 481 AAGTTAGAGGAGGCTGCTCACTATTGTTGTGAATTCAAATCTACTTACAGGCAAGAG 540
 Db |||||
 QY 481 AAGTTAGAGGAGGCTGCTCACTATTGTTGTGAATTCAAATCTACTTACAGGCAAGAG 540
 Db |||||
 QY 541 CCTGTGAGATGCCAGGGTATCATTGTTGACCGCAAACTCGATGTAACCAATCACAA 600
 Db |||||
 QY 541 CCTGTGAGATGCCAGGGTATCATTGTTGACCGCAAACTCGATGTAACCAATCACAA 600
 Db |||||
 QY 601 AAGGATTACACTTCCGTTGAGCAGTGTGAATTCATTTGCAAGCAAACTGCTGCTGCC 660
 Db |||||
 QY 601 AAGGATTACACTTCCGTTGAGCAGTGTGAATTCATTTGCAAGCAAACTGCTGCTGCC 660
 Db |||||
 QY 661 TG 662
 Db |||||
 QY 661 TG 662

RESULT 5
 ABZ12084
 ID ABZ12084 standard; DNA; 660 BP.
 XX AC
 XX ABZ12084;
 XX
 XX 22-JAN-2003 (first entry)
 XX
 XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 51.
 DE
 XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 XX chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunscreen; ds.
 XX
 XX Millepora sp.
 OS
 XX WO200270703-A2.
 FN
 XX 12-SEP-2002.
 PD
 XX 01-MAR-2002; 2002WO-GB000928.
 PF
 XX 02-MAR-2001; 2001US-0273227P.
 PR
 XX 21-MAR-2001; 2001AU-00003874.
 PR
 XX 15-OCT-2001; 2001US-0329816P.
 PR
 XX (NUFA-) NUFARM LTD.
 PA (UYQU) UNIV QUEENSLAND.
 PA (JONE/) JONES E L.
 XX
 XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
 PI Hough-Guldborg IO, Prescott M;
 PI
 XX WPI; 2002-740765/80.
 DR
 XX Novel color-facilitating molecule for producing a biomatrix, has a
 XX polypeptide which alone/along with molecules imparts altered visual
 PT characteristics to cells in the absence of excitation by extraneous non-
 PT white light.
 PT
 XX Claim 6; Page 326-327; 510pp; English.
 PS
 XX The invention relates to an isolated colour-facilitating molecule (CFM)
 CC comprising a polypeptide which, in a cell, alone or together with one or
 CC more other molecules imparts an altered visual characteristic to the cell
 CC when visualised by a human eye in the absence of excitation by extraneous
 CC non-white light or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant
 CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development of
 CC new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
 CC sequences
 XX
 XX Sequence 660 BP; 192 A; 146 C; 163 G; 159 T; 0 U; 0 Other;
 SQ
 Query Match 77.1%; Score 648.8; DB 6; Length 660;
 Best Local Similarity 98.9%; Pred. No. 1.3e-144;
 Matches 653; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTTCAGGCAGCGTCAATGGACAC 60
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTTCAGGCAGCGTCAATGGACAC 60
 Qy 61 TACTTTGAGGTCGAGAGCGGATGGAAGAAAGAGCCTTACGAGGGGAGCAGACGTAAGG 120
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 TACTTTGAGGTCGAGAGCGGATGGAAGAAAGAGCCTTACGAGGGGAGCAGACGTAAGG 120
 Qy 121 CTGGCTGTCAACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGT 180
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 CTGACTGTCAACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTCA 180
 Qy 181 CAGTACGGAAGATACATTCACCAAGTACCTCCCTGAAGACATCCCTGACTATGTAAAGCAG 240
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 CAGTACGGAAGATACATTCACCAAGTACCTCCCTGAAGACATCCCTGACTATGTAAAGCAG 240
 Qy 241 TCATTCCCGGGAGATATACATGGGAGGATCATGAACCTTTGAAGATGGTGCAGTGTGT 300
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 TCATTCCCGGGAGGATATACATGGGAGGATCATGAACCTTTGAAGATGGTGCAGTGTGT 300
 Qy 301 ACTGTCAAGCAATGATTCAGCATCCAGGCAACTGTTTTCATCTACCATGTCAAGTTCTCT 360
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 ACTGTCAAGCAATGATTCAGCATCCAGGCAACTGTTTTCATCTACCATGTCAAGTTCTCT 360
 Qy 361 GGTGTGAATCTTCTCCCAATGGACCTGTTATGCAAGAAAGACACAGGGCTGGGAACCC 420
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 GGTGTGAATCTTCTCCCAATGGACCTGTTATGCAAGAAAGACACAGGGCTGGGAACCC 420
 Qy 421 AACACTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGGAAACAACTTTATGCTCTG 480
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 421 AACACTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGGAAACAACTTTATGCTCTG 480
 Qy 481 AAGTTAGAAGGAGGTGTCCTATTGTTGTAATTCAAATCTTCTTCAAGGCAAGAGAG 540
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 481 AAGTTAGAAGGAGGTGTCCTATTGTTGTAATTCAAATCTTCTTCAAGGCAAGAGAG 540
 Qy 541 CCTGTGAAGATGCCAGGGTATCATTGTTGACCCGAACTGGATGTAACCAATCACAC 600
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 541 CCTGTGAAGATGCCAGGGTATCATTGTTGACCCGAACTGGATGTAACCAATCACAC 600
 Qy 601 AAGGATTACACTTCGTTGAGCAGTGTGAAATTTCCATTGACGCAAACTGTTGTCGCC 660
 XX ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 601 AAGGATTACACTTCGTTGAGCAGTGTGAAATTTCCATTGACGCAAACTGTTGTCGCC 660
 RESULT 6
 ABZ12085
 ID ABZ12085 standard; DNA; 660 BP.
 XX AC
 XX ABZ12085;
 XX 22-JAN-2003 (first entry)
 DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 53.
 XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunscreen; ds.
 XX
 XX Millepora sp.
 OS
 XX WO200270703-A2.
 FN
 XX 12-SEP-2002.
 PD
 XX 01-MAR-2002; 2002WO-GB000928.
 PF
 XX 02-MAR-2001; 2001US-0273227P.
 PR
 XX 21-MAR-2001; 2001AU-00003874.
 PR
 XX 15-OCT-2001; 2001US-0329816P.
 PR
 XX (NUFA-) NUFARM LTD.
 PA (UYQU) UNIV QUEENSLAND.
 PA (JONE/) JONES E L.

XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
 PI Hoegh-Guldberg IO, Prescott M;
 XX WPI; 2002-740765/80.
 XX Novel color-facilitating molecule for producing a biomatrix, has a
 PT polypeptide which alone/along with molecules imparts altered visual
 PT characteristics to cells in the absence of excitation by extraneous non-
 XX white light.
 XX Claim 6; Page 329; 510pp; English.
 XX The invention relates to an isolated colour-facilitating molecule (CFM)
 CC comprising a polypeptide which, in a cell, alone or together with one or
 CC more other molecules imparts an altered visual characteristic to the cell
 CC when visualised by a human eye in the absence of excitation by extraneous
 CC non-white light or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant
 CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development
 CC of new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 XX Sequence 660 BP; 192 A; 147 C; 163 G; 158 T; 0 U; 0 Other;
 SQ
 Query Match 77.0%; Score 647.2; DB 6; Length 660;
 Best Local Similarity 98.8%; Pred. No. 3.1e-144;
 Matches 652; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 TCCGTTATCGTAAACAGATGACCTACAAAGTTTATATGTCAGGCACGGTCAATGGAC 60
 DB 1 TCCGTTATCGTAAACAGATGACCTACAAAGTTTATATGTCAGGCACGGTCAATGGAC 60
 QY 61 TACTTTGAGTTCGAAGGCGATGGAAGGAAAGCCTTACGAGGGGAGCAGACGGTAAGG 120
 DB 61 TACTTTGAGTTCGAAGGCGATGGAAGGAAAGCCTTACGAGGGGAGCAGACGGTAAGG 120
 QY 121 CTGGCTGTACAGAGGGCGACCTCTGCCATTTGCTTGGGATATTTATACCAAGTGT 180
 DB 121 CTGACTGTACCAAGGGCGACCTCTGCCATTTGCTTGGGATATTTATACCAAGTGT 180
 QY 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGATATGTAAGCAG 240
 DB 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGATATGTAAGCAG 240
 QY 241 TCATTCCGGGGAGATATACATGAGGAGGATCATGAACCTTTGAAGATGTCAGTGTGT 300
 DB 241 TCATTCCGGGGAGGATATACATGAGGAGGATCATGAACCTTTGAAGATGTCAGTGTGT 300
 QY 301 ACTGTCAGCATGATTCAGCATTCAGGCAATCTGTTTCATCTACCATGTCAATTTCTCT 360
 DB 301 ACTGTCAGCATGATTCAGCATTCAGGCAATCTGTTTCATCTACCATGTCAATTTCTCT 360
 QY 361 GGTTTGAACCTTCTCCCAATGGACCTTATGAGAGGAGACACAGGGCTGGGAACCC 420
 DB 361 GGTTTGAACCTTCTCCCAATGGACCTTATGAGAGGAGACACAGGGCTGGGAACCC 420
 QY 421 AACACTGAGGCTCTCTTTGACAGAGATGGAATGCTGATAGGAACAACTTTATGGCTGTG 480
 DB 421 AACACTGAGGCTCTCTTTGACAGAGATGGAATGCTGATAGGAACAACTTTATGGCTGTG 480

QY 481 AAGTTAGAGGAGGTGGTCACTATTTGTTGTAATTCAAATCTACTTCAAGGCAAGAAG 540
 DB |||||
 QY 481 AAGTTAGAGGAGGTGGTCACTATTTGTTGTAATTCAAATCTACTTCAAGGCAAGAAG 540
 DB |||||
 QY 541 CCTGTGAAGATGCCAGGATATCACTATGTTGACCCCAAACTGGATGAACCAATCAACA 600
 DB |||||
 QY 541 CCTGTGAAGATGCCAGGATATCACTATGTTGACCCCAAACTGGATGAACCAATCAACA 600
 DB |||||
 QY 601 AAGGATTACACTTCGTTGACGAGTGAATTTCCATTGACGCAAACTGTGTGCGCC 660
 DB |||||
 QY 601 AAGGATTACACTTCGTTGACGAGTGAATTTCCATTGACGCAAACTGTGTGCGCC 660
 DB |||||

RESULT 7
 ABZ12074
 ID ABZ12074 standard; DNA; 660 BP.
 XX
 AC ABZ12074;
 XX
 DT 07-AUG-2003 (revised)
 DT 22-JAN-2003 (first entry)
 XX
 DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 31.
 XX
 KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunscreens; ds.
 XX
 OS Caulastrea sp.
 XX
 PN WO200270703-A2.
 XX
 PD 12-SEP-2002.
 XX
 PF 01-MAR-2002; 2002WO-GB000928.
 XX
 PR 02-MAR-2001; 2001US-0273227P.
 PR 21-MAR-2001; 2001AU-0000387A.
 PR 15-OCT-2001; 2001US-0329816P.
 XX
 PA (NUPA-) NUFARM LTD.
 PA (UYOU) UNIV QUEENSLAND.
 PA (JONE/) JONES E L.
 XX
 PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
 PI Hoegh-Guldberg IO, Prescott M;
 XX
 WPI; 2002-740765/80.
 XX
 Novel color-facilitating molecule for producing a biomatrix, has a
 PT polypeptide which alone/along with molecules imparts altered visual
 PT characteristics to cells in the absence of excitation by extraneous non-
 XX white light.
 XX Claim 6; Page 297-298; 510pp; English.
 XX The invention relates to an isolated colour-facilitating molecule (CFM)
 CC comprising a polypeptide which, in a cell, alone or together with one or
 CC more other molecules imparts an altered visual characteristic to the cell
 CC when visualised by a human eye in the absence of excitation by extraneous
 CC non-white light or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant
 CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development
 CC of new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their

CC marketability, CFMs embedded in a gel matrix improve image quality in
CC situations of distorted light spectra (biomatrix). The first all-protein
CC chromophore to be isolated was Green Fluorescent protein (GFP). The
CC sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
CC sequences. (Updated on 07-AUG-2003 to correct OS field.)
XX
SQ Sequence 660 BP; 192 A; 147 C; 163 G; 158 T; 0 U; 0 Other;

Query Match 77.0%; Score 647.2; DB 6; Length 660;
Best Local Similarity 98.8%; Pred. No. 3.1e-144;
Matches 652; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 TCCGTTATCGTAAACAGATGACCTACAAAGTTATATGTCTAGGACGCTCAATGGACAC 60
DB |||||
QY 1 TCCGTTATCGTAAACAGATGACCTACAAAGTTATATGTCTAGGACGCTCAATGGACAC 60
DB |||||
QY 61 TACTTTGAGGTCGAAGGCGATGGAAGGAAAGCCCTTACGAGGGGAGCAGACGGTAAAG 120
DB |||||
QY 61 TACTTTGAGGTCGAAGGCGATGGAAGGAAAGCCCTTACGAGGGGAGCAGACGGTAAAG 120
DB |||||
QY 121 CTGGCTGTCCAGAGGGGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTGT 180
DB |||||
QY 121 CTGACTGTCCAGAGGGGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTCA 180
DB |||||
QY 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAGCAG 240
DB |||||
QY 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAGCAG 240
DB |||||
QY 241 TCATTCCCGGGAGATATACATGGAGAGGATCATGAACTTTGAAAGTGGTCAAGTGT 300
DB |||||
QY 241 TCATTCCCGGGAGGATATACATGGAGAGGATCATGAACTTTGAAAGTGGTCAAGTGT 300
DB |||||
QY 301 ACTGTCCAGCATGATCCAGCATCCAGGCACTGTTTCATCTACCATGTCAGTTCCT 360
DB |||||
QY 301 ACTGTCCAGCATGATCCAGGCACTGTTTCATCTACCATGTCAGTTCCT 360
DB |||||
QY 361 GGTTTGAACTTCTCCCAATGGAACCTGTATGCAAGAGACACAGGCTGGGAACC 420
DB |||||
QY 361 GGTTTGAACTTCTCCCAATGGAACCTGTATGCAAGAGACACAGGCTGGGAACC 420
DB |||||
QY 421 AACACTGAGCGTCTCTGTCAGAGATGGAATGCTGATAGGAAACAACCTTTATGGCTCTG 480
DB |||||
QY 481 AAGTTAGAGAGGCTGTGCTACATTTGCTGAATTCATTAATCTACTTACAGGCAAGAG 540
DB |||||
QY 481 AAGTTAGAGAGGCTGTGCTACATTTGCTGAATTCATTAATCTACTTACAGGCAAGAG 540
DB |||||
QY 541 CCTGTGAAGATCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAC 600
DB |||||
QY 541 CCTGTGAAGATCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAC 600
DB |||||
QY 601 AAGGATTACACTTCCGTTGAGCAGTGTGAAATTCATTGCGCGMAAACCCTGTGTCGCC 660
DB |||||
QY 601 AAGGATTACACTTCCGTTGAGCAGCGTGAAATTCATTGCGCGMAAACCCTGTGTCGCC 660
DB |||||

RESULT 8
ID ABZ12069 standard; DNA; 660 BP.
XX
AC ABZ12069;
XX
XX 22-JAN-2003 (first entry)
XX
DE Colour Facilitating molecule (CFM) related sequence #SBQ ID 21.
XX
KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
XX UV sink; sunscreen; ds.
XX
OS Acropora aspera.

XX WO200270703-A2.
XX
XX 12-SEP-2002.
XX
XX 01-MAR-2002; 2002WO-GB000928.
XX
XX 02-MAR-2001; 2001US-0273227P.
XX 21-MAR-2001; 2001AU-00003874.
XX 15-OCT-2001; 2001US-0329816P.
XX
XX (NUFA-) NUFARM LTD.
XX (UYQU) UNIV QUEENSLAND.
XX (JONE/) JONES E L.
XX
XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
XX Hough-Guldberg IO, Prescott M;
XX
XX WPI; 2002-740765/80.
XX
XX Novel color-facilitating molecule for producing a biomatrix, has a
XX polypeptide which alone/along with molecules imparts altered visual
XX characteristics to cells in the absence of excitation by extraneous non-
XX white light.
XX
XX Claim 6; Page 285-286; 510pp; English.
XX
XX The invention relates to an isolated colour-facilitating molecule (CFM)
XX comprising a polypeptide which, in a cell, alone or together with one or
XX more other molecules imparts an altered visual characteristic to the cell
XX when visualised by a human eye in the absence of excitation by extraneous
XX non-white light or particle emission. CFMs are useful for producing a
XX transgenic animal which exhibits a novel colour e.g. sheep with blue or
XX red coloured fleeces. They are useful for producing coloured plant
XX extracts, e.g. flavouring, beverage or juice or colouring agent. Other
XX uses include transducing or intensifying an image, providing additional
XX light for growing phototropic organisms e.g. algae and/or corals, for
XX coating materials that experience UV damage e.g. plastics and car
XX upholstery. CFMs are useful in the flower industry, in the development of
XX new varieties of flowering plants. Other contemplated uses include,
XX expression markers, general reporter molecules, photon traps, UV sinks or
XX in sunscreens. CFMs modify visible colour in edible and/or ornamental
XX fungal species, and in fruits and vegetables to enhance their
XX marketability. CFMs embedded in a gel matrix improve image quality in
XX situations of distorted light spectra (biomatrix). The first all-protein
XX chromophore to be isolated was Green Fluorescent protein (GFP). The
XX sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
XX sequences
XX
XX Sequence 660 BP; 193 A; 146 C; 163 G; 158 T; 0 U; 0 Other;
SQ

Query Match 77.0%; Score 647.2; DB 6; Length 660;
Best Local Similarity 98.8%; Pred. No. 3.1e-144;
Matches 652; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 TCCGTTATCGTAAACAGATGACCTACAAAGTTATATGTCTAGGACGCTCAATGGACAC 60
DB |||||
QY 1 TCCGTTATCGTAAACAGATGACCTACAAAGTTATATGTCTAGGACGCTCAATGGACAC 60
DB |||||
QY 61 TACTTTGAGGTCGAAGGCGATGGAAGGAAAGCCCTTACGAGGGGAGCAGACGGTAAAG 120
DB |||||
QY 61 TACTTTGAGGTCGAAGGCGATGGAAGGAAAGCCCTTACGAGGGGAGCAGACGGTAAAG 120
DB |||||
QY 121 CTGGCTGTCCAGAGGGGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTGT 180
DB |||||
QY 121 CTGACTGTCCAGAGGGGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTCA 180
DB |||||
QY 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAGCAG 240
DB |||||
QY 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAGCAG 240
DB |||||
QY 241 TCATTCCCGGGAGATATACATGGAGAGGATCATGAACTTTGAAAGTGGTCAAGTGT 300
DB |||||

Db 241 TCATTCCCGAGGGATATACATGGAGAGGATCATGAATTTGAAGATGGTGCAGTGTGT 300
QY 301 ACTGTGAGCAATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT 360
Db 301 ACTGTGAGCAATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT 360
QY 361 GGTTTGAATCTTCTCCCAATGGACCTGTTATGCAAGAAGACACAGGCTGGGAACCC 420
Db 361 GGTTTGAATCTTCTCCCAATGGACCTGTTATGCAAGAAGACACAGGCTGGGAACCC 420
QY 421 AACCTGAGCGTCTCTTTGCAGCAGATGGAATGCTGATAGGAAACAATTTATGCTCTG 480
Db 421 AACCTGAGCGTCTCTTTGCAGCAGATGGAATGCTGATAGGAAACAATTTATGCTCTG 480
QY 481 AAGTTAGAGAGGTGGTGCATATTTGTGTGAATTCAAATCTACATTCACAGGCAAGGAAG 540
Db 481 AAGTTAGAGAGGTGGTGCATATTTGTGTGAATTCAAATCTACATTCACAGGCAAGGAAG 540
QY 541 CCTGTGAAGATGCCAGGGTATCATTATGTTGACCGCAAACTGGATGAACCAATCAACAAC 600
Db 541 CCTGTGAAGATGCCAGGGTATCATTATGTTGACCGCAAACTGGATGAACCAATCAACAAC 600
QY 601 AAGGATTACACTTCGTTGAGCAGTGTGAATTTCCATTGACGCAAACTGGTGTGCGCC 660
Db 601 AAGGATTACACTTCGTTGAGCAGTGTGAATTTCCATTGACGCAAACTGGTGTGCGCC 660

RESULT 9

ABZ12156
ID ABZ12156 standard; DNA; 669 BP.
AC ABZ12156;
XX
DT 22-JAN-2003 (first entry)
DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 211.
XX
KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
KW UV sink; sunscreen; ds.
XX

Discosoma sp.

XX
XX WO200270703-A2.
XX

12-SEP-2002.

XX
XX 01-MAR-2002; 2002WO-GB000928.
XXXX 02-MAR-2001; 2001US-0273227P.
PR 21-MAR-2001; 2001AU-0000387A.
PR 15-OCT-2001; 2001US-0329816P.XX
XX (NUFA-) NUFARM LTD.
PA (UYOU) UNIV QUEENSLAND.
PA (JONE/) JONES E L.
XX

XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
PI Hoegh-Guldberg IO, Prescott M;
XX WPI; 2002-740765/80.
XX

XX Novel color-facilitating molecule for producing a biomatrix, has a
PT polypeptide which alone/along with molecules imparts altered visual
PT characteristics to cells in the absence of excitation by extraneous non-
PT white light.
XX

XX Example 12; Page 482-483; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)
CC comprising a polypeptide which, in a cell, alone or together with one or
CC more other molecules imparts an altered visual characteristic to the cell
CC

CC when visualised by a human eye in the absence of excitation by extraneous
CC non-white light or particle emission. CFMs are useful for producing a
CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
CC red coloured fleece. They are useful for producing coloured plant
CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
CC uses include transducing or intensifying an image, providing additional
CC light for growing phototropic organisms e.g. algae and/or corals, for
CC coating materials that experience UV damage e.g. plastics and car.
CC upholstery. CFMs are useful in the flower industry, in the development of
CC new varieties of flowering plants. Other contemplated uses include,
CC expression markers, general reporter molecules, photon traps, UV sinks or
CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
CC fungal species, and in fruits and vegetables to enhance their
CC marketability. CFMs embedded in a gel matrix improve image quality in
CC situations of distorted light spectra (biomatrix). The first all-protein
CC chromophore to be isolated was Green Fluorescent protein (GFP). The
CC sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
CC sequences
XX

SQ Sequence 669 BP; 195 A; 148 C; 167 G; 159 T; 0 U; 0 Other;

Query Match 77.0%; Score 647.2; DB 6; Length 669;

Best Local Similarity 98.8%; Pred. No. 3.1e-144;

Matches 652; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 TCCGTTATCGCTAAACAGATGACCTTACAAAGTTTATATGTCAGGCACGCTCAATGCACAC 60

Db 4 TCCGTTATCGCTAAACAGATGACCTTACAAAGTTTATATGTCAGGCACGCTCAATGCACAC 63

QY 61 TACTTTGAGTTCGAAGGCGATGGAAGAAAGCCTTACGAGGGGGAGCAGACGGTAAGG 120

Db 64 TACTTTGAGTTCGAAGGCGATGGAAGAAAGCCTTACGAGGGGGAGCAGACGGTAAGG 123

QY 121 CTGCTGTCCACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTGT 180

Db 124 CTGCTGTCCACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTGT 183

QY 181 CAGTACGGAAGCATACATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAAGCAG 240

Db 184 CAGTACGGAAGCATACATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAAGCAG 243

QY 241 TCATTCCCGGGGAGATATACATGGGAGAGGATCATGAATTTGAAGATGGTGCAGTGTGT 300

Db 244 TCATTCCCGGGGAGATATACATGGGAGAGGATCATGAATTTGAAGATGGTGCAGTGTGT 303

QY 301 ACTGTGAGCATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT 360

Db 304 ACTGTGAGCATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT 363

QY 361 GGTTTGAATCTTCTCCCAATGGACCTGTTATGCAAGAAGAGACACAGGGCTGGGAACCC 420

Db 364 GGTTTGAATCTTCTCCCAATGGACCTGTTATGCAAGAAGAGACACAGGGCTGGGAACCC 423

QY 421 AACACTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGGAAACAATTTATGCTCTG 480

Db 424 AACACTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGGAAACAATTTATGCTCTG 483

QY 481 AAGTTAGAAGGAGTGTGCTACTATTTGTGTAATTCAAATCTACTTCAAGGCAAGGAAG 540

Db 484 AAGTTAGAAGGAGTGTGCTACTATTTGTGTAATTCAAATCTACTTCAAGGCAAGGAAG 543

QY 541 CCTGTGAAGATGCCAGGGTATCATTATGTTGACCGCAAACTGGATGTAAACCAATCAACAAC 600

Db 544 CCTGTGAAGATGCCAGGGTATCATTATGTTGACCGCAAACTGGATGTAAACCAATCAACAAC 603

QY 601 AAGGATTACACTTCGTTGAGCAGTGTGAATTTCCATTGACGCAAACTGGTGTGCGCC 660

Db 604 AAGGATTACACTTCGTTGAGCAGTGTGAATTTCCATTGACGCAAACTGGTGTGCGCC 663

RESULT 10

ABZ12072

ID ABZ12072 standard; DNA; 660 BP.

XX ABZ12072;
 XX 07-AUG-2003 (revised)
 XX 22-JAN-2003 (first entry)
 XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 27.
 XX
 XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunscreen; ds.
 XX
 XX Acanthastrea sp.
 XX
 XX WO200270703-A2.
 XX
 XX 12-SEP-2002.
 XX
 XX 01-MAR-2002; 2002WO-GB000928.
 XX
 XX 02-MAR-2001; 2001US-0273227P.
 XX 21-MAR-2001; 2001AU-00003874.
 XX 15-OCT-2001; 2001US-0329816P.
 XX
 XX (NUFA-) NUFARM LTD.
 XX (UYQU) UNIV QUEENSLAND.
 XX (JONE/) JONES E L.
 XX
 XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
 XX Hoegh-Guldberg IO, Prescott M;
 XX
 XX WPI; 2002-740765/80.
 XX
 XX Novel color-facilitating molecule for producing a biomatrix, has a
 PT polypeptide which alone/along with molecules imparts altered visual
 PT characteristics to cells in the absence of excitation by extraneous non-
 PT white light.
 XX
 XX Claim 6; Page 292-293; 510pp; English.
 XX
 XX The invention relates to an isolated colour-facilitating molecule (CFM)
 CC comprising a polypeptide which, in a cell, alone or together with one or
 CC more other molecules imparts an altered visual characteristic to the cell
 CC when visualised by a human eye in the absence of excitation by extraneous
 CC non-white light or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant
 CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development of
 CC new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
 CC sequences. (Updated on 07-AUG-2003 to correct OS field.)
 XX
 XX Sequence 660 BP; 191 A; 147 C; 164 G; 158 T; 0 U; 0 Other;
 SQ
 Query Match 76.8%; Score 645.6; DB 6; Length 660;
 Best Local Similarity 98.6%; Pred. No. 7.4e-144;
 Matches 651; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 1 TCGTTATCGCTAAACAGATGACCTCAAGTTTATATGTCAGGCGGTCATGACAC 60
 1 TCGTTATCGCTAAACAGATGACCTCAAGTTTATATGTCAGGCGGTCATGACAC 60

QY 61 TACTTTGAGGTCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGACGGTAAGG 120
 DB 61 TACTTTGAGGTCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGACGGTAAGG 120
 QY 121 CTGGCTGTCAACAAGGCGGACCTCTGCCATTTGCTGGGATATTTTATCACCACAGTGT 180
 DB 121 CTGACTGTCAACAAGGCGGACCTCTGCCATTTGCTGGGATATTTTATCACCACAGTCA 180
 QY 181 CAGTACGAAAGCATACCAATTCACCAAGTACCTCCTGAAGACATCCCTGACTATGTAAGCG 240
 DB 181 CAGTACGAAAGCATACCAATTCACCAAGTACCTCCTGAAGACATCCCTGACTATGTAAGCG 240
 QY 241 TCATTTCCCGGGGAGATATACATATGGGAGAGGATCATGAACCTTTGAAGATGGTGCAGTGT 300
 DB 241 TCATTTCCCGGGGAGATATACATATGGGAGAGGATCATGAACCTTTGAAGATGGTGCAGTGT 300
 QY 301 ACTGTCAGCAATGATTCAGGATCCAGGCACTGTTTCACTACCATGTCAAGTTCTCT 360
 DB 301 ACTGTCAGCAATGATTCAGGATCCAGGCACTGTTTCACTACCATGTCAAGTTCTCT 360
 QY 361 GGTTTGAACCTTCTCCCAATGGACCTGTTATGCAGAGAGACACACAGGCTGGGAACCC 420
 DB 361 GGTTTGAACCTTCTCCCAATGGACCTGTTATGCAGAGAGAGACACAGGCTGGGAACCC 420
 QY 421 AACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACCTTTATGGCTCTG 480
 DB 421 AACACTGAGCGTCTCTTTGCACGAGATGGAATGCTGATAGGAAACAACCTTTATGGCTCTG 480
 QY 481 AAGTTAGAAGGAGGTGGTCACTATTTGTGTAATTCAAATCTACTTACAGGCAAGAG 540
 DB 481 AAGTTAGAAGGAGGTGGTCACTATTTGTGTAATTCAAATCTACTTACAGGCAAGAG 540
 QY 541 CCTGTGAGAGTCCAGGGTATCCTATGTTGACCGCAACTGGATGTACCAATCAACAAC 600
 DB 541 CCTGTGAGAGTCCAGGGTATCCTATGTTGACCGCAACTGGATGTACCAATCAACAAC 600
 QY 601 AAGGATTACACTTCGCTGGAGCAGTGTGAAATTTCCATTGACGCAAAACCTGTGTGCGCC 660
 DB 601 AAGGATTACACTTCGCTGGAGCAGTGTGAAATTTCCATTGACGCAAAACCTGTGTGCGCC 660
 RESULT 11
 ABZ12155
 ID ABZ12155 standard; DNA; 669 BP.
 XX
 XX ABZ12155;
 XX
 XX 22-JAN-2003 (first entry)
 DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 209.
 XX
 XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunscreen; ds.
 XX
 XX Acropora sp.
 XX
 XX WO200270703-A2.
 XX
 XX 12-SEP-2002.
 XX
 XX 01-MAR-2002; 2002WO-GB000928.
 XX
 XX 02-MAR-2001; 2001US-0273227P.
 XX 21-MAR-2001; 2001AU-00003874.
 XX 15-OCT-2001; 2001US-0329816P.
 XX
 XX (NUFA-) NUFARM LTD.
 XX (UYQU) UNIV QUEENSLAND.
 XX (JONE/) JONES E L.
 XX
 XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;

PI Hoegh-Guldberg IO, Prescott M;
XX WPI; 2002-740765/80.
DR Novel color-facilitating molecule for producing a biomatrix, has a
XX polypeptide which alone/along with molecules imparts altered visual
PT characteristics to cells in the absence of excitation by extraneous non-
PT white light.
XX Example 11; Page 481; 510pp; English.
XX The invention relates to an isolated colour-facilitating molecule (CFM)
CC comprising a polypeptide which, in a cell, alone or together with one or
CC more other molecules imparts an altered visual characteristic to the cell
CC when visualised by a human eye in the absence of excitation by extraneous
CC non-white light or particle emission. CFMs are useful for producing a
CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
CC red coloured fleece. They are useful for producing coloured plant
CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
CC uses include transducing or intensifying an image, providing additional
CC light for growing phototropic organisms e.g. algae and/or corals, for
CC coating materials that experience UV damage e.g. plastics and car
CC upholstery. CFMs are useful in the flower industry, in the development of
CC new varieties of flowering plants. Other contemplated uses include
CC expression markers, general reporter molecules, photon traps, UV sinks or
CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
CC fungal species, and in fruits and vegetables to enhance their
CC marketability. CFMs embedded in a gel matrix improve image quality in
CC situations of distorted light spectra (biomatrix). The first all-protein
CC chromophore to be isolated was Green Fluorescent protein (GFP). The
CC sequences given in records AB212068-AB212195 represent CFM related DNA
XX sequences
SQ Sequence 669 BP; 195 A; 148 C; 166 G; 160 T; 0 U; 0 Other;
Query Match 76.4%; Score 642.4; DB 6; Length 669;
Best Local Similarity 98.3%; Pred No. 4.3e-143;
Matches 549; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 1 TCCGTTATCGTTAAACAGATGACCTTACAAAGTTTATATGTCAGGCAGCGTCAATGGACAC 60
DB 4 TCCGTTATCGTTAAACAGATGACCTTACAAAGTTTATATGTCAGGCAGCGTCAATGGACAC 63
QY 61 TACTTTGAGTGGTGAAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGCGTAAAG 120
DB 64 TACTTTGAGTGGTGAAGGCGATGGAAGAAAGCCCTTACGAGGGGAGCAGCGTAAAG 123
QY 121 CTGGCTGTCAACAAAGGGGACCTCTGCCATTTCCTTGGGATATTTTATCACCACAGTGT 180
DB 124 CTCAGTGTCAACAAAGGGTGGACCTCTGCCATTTCCTTGGGATATTTTATCACCACAGTCA 183
QY 181 CAGTACGGAAGCATACCATTCACCAAGTACCTTGAAGACATCCCTTGACTATGTAAGCAG 240
DB 184 CAGTACGGAAGCATACCATTCACCAAGTACCTTGAAGACATCCCGGACTATGTAAGCAG 243
QY 241 TCATTCCCGGGAGATATACATGGAGAGATCATGAACCTTTGAAGTGGTGCAGTGTGT 300
DB 244 TCATTCCCGGGAGGATATACATGGAGAGGATCATGAACCTTTGAAGTGGTGCAGTGTGT 303
QY 301 ACTGTGACATGATTTCCAGATCCAGGCATCCAGGCACTGTTTCATCTACCATGTCAAGTTCTCT 360
DB 304 ACTGTGACATGATTTCCAGATCCAGGCATCCAGGCACTGTTTCATCTACCATGTCAAGTTCTCT 363
QY 361 GGTTTGAACCTTCTCCCAATGGACCTGTTATGAGAGAAAGACACAGGGCTGGGAACC 420
DB 364 GGTTTGAACCTTCTCCCAATGGACCTGTTATGAGAGAAAGACACAGGGCTGGGAACC 423
QY 421 AACACTGAGCTCTCTTTGACAGATGGATGCTGATAGAAACAACTTTATGGCTCTG 480
DB 424 AACACTGAGCTCTCTTTGACAGATGGATGCTGATAGAAACAACTTTATGGCTCTG 483
QY 481 AAGTTAGAAGAGGTGGTGCATATTGTGTGAATTCAAATCTACTTACAGGCAAGAAAG 540
DB 484 AAGTTAGAAGAGGTGGTGCATATTGTGTGAATTCAAATCTACTTACAGGCAAGAAAG 543
QY 541 CCTGTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCTACTTACAGGCAAGAAAG 600
DB 544 CCTGTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCTACTTACAGGCAAGAAAG 603
QY 601 AAGGATTACACTTCCGTTGAGCAGTGTGAAATTTCCATTTGACGCAAACTGGTGTGGTGGCC 660
DB 604 AAGGATTACACTTCCGTTGAGCAGTGTGAAATTTCCATTTGACGCAAACTGGTGTGGTGGCC 663
RESULT 12
AB212161
ID AB212161 standard; DNA; 669 BP.
XX
AC AB212161;
DT 22-JAN-2003 (first entry)
XX
DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 221.
XX
KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
KW UV sink; sunsreen; ds.
XX
OS Tubastrea sp.
XX
PN WO200270703-A2.
XX
PD 12-SEP-2002.
XX
PF 01-MAR-2002; 2002WO-GB000928.
XX
PR 02-MAR-2001; 2001US-0273227P.
PR 21-MAR-2001; 2001AU-00003874.
PR 15-OCT-2001; 2001US-0329816P.
XX
PA (NUFA-) NUFARM LTD.
PA (UYOU) UNIV QUEENSLAND.
XX (JONE/) JONES E L.
PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
PI Hoegh-Guldberg IO, Prescott M;
XX
DR WPI; 2002-740765/80.
XX
PT Novel color-facilitating molecule for producing a biomatrix, has a
PT polypeptide which alone/along with molecules imparts altered visual
PT characteristics to cells in the absence of excitation by extraneous non-
PT white light.
XX
XX Example 18; Page 490; 510pp; English.
XX The invention relates to an isolated colour-facilitating molecule (CFM)
CC comprising a polypeptide which, in a cell, alone or together with one or
CC more other molecules imparts an altered visual characteristic to the cell
CC when visualised by a human eye in the absence of excitation by extraneous
CC non-white light or particle emission. CFMs are useful for producing a
CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
CC red coloured fleece. They are useful for producing coloured plant
CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
CC uses include transducing or intensifying an image, providing additional
CC light for growing phototropic organisms e.g. algae and/or corals, for
CC coating materials that experience UV damage e.g. plastics and car
CC upholstery. CFMs are useful in the flower industry, in the development of
CC new varieties of flowering plants. Other contemplated uses include
CC expression markers, general reporter molecules, photon traps, UV sinks or
CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
CC fungal species, and in fruits and vegetables to enhance their
CC marketability. CFMs embedded in a gel matrix improve image quality in
CC situations of distorted light spectra (biomatrix). The first all-protein
CC chromophore to be isolated was Green Fluorescent protein (GFP). The

CC	sequences given in records ABZ12068--ABZ12195 represent CFM related DNA
CC	sequences
XX	
SQ	Sequence 669 BP; 189 A; 147 C; 170 G; 163 T; 0 U; 0 Other;
Query Match 76.2%; Score 640.8; DB 6; Length 669;	
Best Local Similarity 98.2%; Pred. No. 1e-142;	
Matches 648; Conservative 0; Mismatches 12; Indels 0; Gaps 0;	
QY	1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGCACGGTCAATGGACAC 60
DB	
QY	4 TCCGTTATCGCTAAACAGATGACCTACAGTTTATATGTCAGGCACGGTCAATGGACAC 63
DB	
QY	61 TACTTTGAGGTGCAAGGCGATGGAAGAAAGGAAAGCCTTACGAGGGGAGCAGACGTAAGG 120
DB	
QY	64 TACTTTGAGGTGCAAGGCGATGGAAGAAAGGAAAGCCTTACGAGGGGAGCAGACGTAAGG 123
DB	
QY	121 CTGGCTGTACCAAGGGCGACCTCTGCGCATTTGCTTGGGATATTTTATCACACAGTGT 180
DB	
QY	124 CTGGCTGTACCAAGGGCGACCTCTGCGCATTTGCTTGGGATATTTTATCACACAGTGT 183
DB	
QY	181 CAGTACGGAAGCATACCATTCACCAAGTACCCTCAAGACATCCCTGACTATGTAAGACAG 240
DB	
QY	184 CAGTACGGAAGCATACCATTCACCAAGTACCCTCAAGACATCCCTGACTATGTAAGACAG 243
DB	
QY	241 TCATTTCCGGGGAGATATACATGGGAGAGATCATGAACCTTTGAAGATGTCAGTGTGT 300
DB	
QY	244 TCATTTCCGGGGAGATTTACATGGGAGAGATCATGAACCTTTGAAGATGTCAGTGTGT 303
DB	
QY	301 ACTGTCAGCAATGATTCAGGATCCAAAGGGAACCTGTTTCATCTACATGTCAAGTTCCT 360
DB	
QY	304 ACTGTCAGCAATGATTCAGGATCCAAAGGGAACCTGTTTCATCTACATGTCAAGTTCCT 363
DB	
QY	361 GGTTTGAACCTTCCCTCCCAATGGACCTGTTATGCAGAAAGACACAGGGCTGGGAACCC 420
DB	
QY	364 GGTTTGAACCTTCCCTCCCAATGGACCTGTTATGCAGAAAGACACAGGGCTGGGAACCC 423
DB	
QY	421 AACACTGAGCGCTCTCTTTGACGAGATGGAATGCTGATAGGAACAACCTTTATGSGCTCTG 480
DB	
QY	424 CACTCTGAGCGTCTCTTTGACGAGACGGAATGCTGATAGGAACAACCTTTATGSGCTCTG 483
DB	
QY	481 AAGTTAGAGAGAGTGTCATTTTGTGTGAATTCAAATCTACTTACAAGGCMAAGAG 540
DB	
QY	484 AAGTTAGAGAGAGCGGTCACTATTTGTGTGAATTCAAACTTACTTACAAGGCMAAGAG 543
DB	
QY	541 CCTGTGAGATGCCAGGGTATCACTATGTCACGCAACCTGGATGAACCAATCACAC 600
DB	
QY	544 CCTGTGAGATGCCAGGGTATCACTATGTCACGCAACCTGGATGAATCAATCACAC 603
DB	
QY	601 AAGGATTACACTTCGGTTGACAGTGTGAATTTCCATTTGCACGCAACCTGTGTGTCGCC 660
DB	
QY	604 AAGGATTACACTTCGGTTGACAGTGTGAATTTCCATTTGCACGCAACCTGTGTGTCGCC 663
DB	
RESULT 13	
ABZ12158	
ID	ABZ12158 standard; DNA; 669 BP.
XX	
AC	ABZ12158;
XX	
DT	22-JAN-2003 (first entry)
XX	
DE	Colour Facilitating molecule (CFM) related sequence #SEQ ID 215.
XX	
KW	Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW	chromophore; biomatix; transgenic animal; colouring agent;
KW	flower industry; expression marker; reporter molecule; photon trap;
KW	UV sink; screen; ds.
OS	Tubastrea sp.
XX	
FN	WO200270703-A2.
XX	

PD	12-SEP-2002.
XX	
PF	01-MAR-2002; 2002WO-GS000928.
XX	
PR	02-MAR-2001; 2001US-0273227P.
PR	21-MAR-2001; 2001AU-0000387A.
PR	15-OCT-2001; 2001US-0329816P.
XX	(NUFA-) NUFARM LTD.
PA	(UYOU) UNIV QUEENSLAND.
PA	(JONE/) JONES E L.
XX	
PI	Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
PI	Hoegh-Guldberg IO, Prescott M;
XX	
DR	WPI; 2002-740765/80.
XX	
PT	Novel color-facilitating molecule for producing a biomatrix, has a
PT	polypeptide which alone/along with molecules imparts altered visual
PT	characteristics to cells in the absence of excitation by extraneous non-
PT	white light.
XX	
PS	Example 18; Page 485-486; 510pp; English.
XX	
CC	The invention relates to an isolated colour-facilitating molecule (CFM)
CC	comprising a polypeptide which, in a cell, alone or together with one or
CC	more other molecules imparts an altered visual characteristic to the cell
CC	when visualised by a human eye in the absence of excitation by extraneous
CC	non-white light or particle emission. CFMs are useful for producing a
CC	transgenic animal which exhibits a novel colour e.g. sheep with blue or
CC	red coloured fleece. They are useful for producing coloured plant
CC	extracts, e.g. flavouring, beverage or juice or colouring agent. Other
CC	uses include transducing or intensifying an image, providing additional
CC	light for growing phototropic organisms e.g. algae and/or corals, for
CC	coating materials that experience UV damage e.g. plastics and car
CC	upholstery. CFMs are useful in the flower industry, in the development of
CC	new varieties of flowering plants. Other contemplated uses include,
CC	expression markers, general reporter molecules, photon traps, UV sinks or
CC	in sunscreens. CFMs modify visible colour in edible and/or ornamental
CC	fungal species, and in fruits and vegetables to enhance their
CC	marketability. CFMs embedded in a gel matrix improve image quality in
CC	situations of distorted light spectra (biomatrix). The first all-protein
CC	chromophore to be isolated was Green Fluorescent protein (GFP). The
CC	sequences given in records AB212068-AB212195 represent CFM related DNA
XX	sequences
XX	
SQ	Sequence 669 BP; 194 A; 148 C; 167 G; 160 T; 0 U; 0 Other;
	Query Match 76.2%; Score 640.8; DB 6; Length 669;
	Best Local Similarity 98.2%; Pred. No. ie-142;
	Matches 648; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
Qy	1 TC CGTTATCGCTAAACAGTACCTCAAGATTATATGTTCAGCAGCGTGCAATGGACAC 60
Db	
	4 TC CGTTATCGCTAAACAGTAGACTCAAGTTTTATATGTTCAGCAGCGTTCAATGGACAC 63
Qy	61 TACTTTTCAGGTTCGAAGGCATGAAAAAGGAAAAGCCTTACGAGGGGGAGCACACGGTAAGG 120
Db	
	64 TACTTTTCAGGTTCGAAGGCATGAAAAAGGAAAAGCCTTACGAGGGGGAGCACACGGTAAG 123
Qy	121 CTGGCTGTCAACAAAGGGCGACCTCTGCCAATTCGCTTGGGATATTTTATACCAACAGTGT 180
Db	
	124 CTCACGTGTCAACAAAGGTGGACCTCTGCCAATTCGCTTGGGATATTTTATACCAACAGTCA 183
Qy	181 CAGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAGCAG 240
Db	
	184 CAGTACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCGGACTATGTAAGCAG 243
Qy	241 TCATTCCCGGGGAGATATACATGGAGAGGATCATGAACCTTGAAGATGGTCAAGTGTGT 300
Db	
	244 TCATTCCCGGGGAGATATACATGGAGAGGATCATGAACCTTGAAGATGGTCAAGTGTGT 303
Qy	301 ACTGTTCAGCATGATTCACGATCCAAAGGCAACTGTTTTCATCTACCATGTCAAGTTCTCT 360

Db 304 ACTGTCAGCAATGACTCCAGCATCCAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT 363
Qy 361 GGTTTGAACCTTCCCTCCCAATGGACCTGTTATGCAGAGAGACACACAGGCTGGGAACCC 420
Db 364 GGTTTGAACCTTCCCTCCCAATGGACCTGTTATGCAGAGAGACACACAGGCTGGGAACCC 423
Qy 421 AACACTGAGCGCTCTCTTTGCAGAGATGGAATGCTGATAGGAACAACTTTATGCTCTG 480
Db 424 AACACTGAGCGCTCTCTTTGCAGAGATGGAATGCTGATAGGAACAACTTTATGCTCTG 483
Qy 481 AAGTTAGAGAGAGGCTGCTATTTGTTGTAATTCATCTCAATCTTACAGGCAAGGAAG 540
Db 484 AAGTTAGAGAGAGGCTGCTATTTGTTGTAATTCATCTCAATCTTACAGGCAAGGAAG 543
Qy 541 CCTGTGAAGATGCCAGGCTGATCACTATGTTGACCGCAAACTGGATGAACCAATCACAAC 600
Db 544 CCTGTGAAGATGCCAGGCTGATCACTATGTTGACCGCAAACTGGATGAACCAATCACAAC 603
Qy 601 AAGGATTACACTCCGTTGAGCAGTGTGAAATTTCCATTTGCAACGCAAACTGTGTGCGCC 660
Db 604 AAGGATTACACTCCGTTGAGCAGTGTGAAATTTCCATTTGCGCGCAAACTGTGTGCGCC 663

RESULT 14

ABZ12083
ID ABZ12083 standard; DNA; 660 BP.
AC
XX
AC ABZ12083;
XX
DT 22-JAN-2003 (first entry)
XX
XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 49.
DE
XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
KW UV sink; sunsreen; ds.
XX
OS Millepura sp.
XX
XX WO200270703-A2.
XX
XX
PD 12-SEP-2002.
XX
XX 01-MAR-2002; 2002WO-GB000928.
XX
XX 02-MAR-2001; 2001US-0273227P.
PR 21-MAR-2001; 2001AU-00003874.
PR 15-OCT-2001; 2001US-0329816P.
XX
XX (NUFA-) NUFARM LTD.
PA (UYOU) UNIV QUEBENSLAND.
PA (JONE/) JONES E L.
XX
PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
PI Hoegh-Guldberg IO, Prescott M;
XX
XX WPI; 2002-740765/80.

XX Novel color-facilitating molecule for producing a biomatrix, has a
XX polypeptide which alone/along with molecules imparts altered visual
XX characteristics to cells in the absence of excitation by extraneous non-
XX white light.
XX
XX Claim 6; Page 324-325; 510pp; English.
XX
XX The invention relates to an isolated colour-facilitating molecule (CFM)
XX comprising a polypeptide which, in a cell, alone or together with one or
XX more other molecules imparts an altered visual characteristic to the cell
XX when visualised by a human eye in the absence of excitation by extraneous
XX non-white light or particle emission. CFMs are useful for producing a
XX transgenic animal which exhibits a novel colour e.g. sheep with blue or

CC red coloured fleece. They are useful for producing coloured plant
CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
CC uses include transducing or intensifying an image, providing additional
CC light for growing phototropic organisms e.g. algae and/or corals, for
CC coating materials that experience UV damage e.g. plastics and car
CC upholstery. CFMs are useful in the flower industry, in the development of
CC new varieties of flowering plants. Other contemplated uses include,
CC expression markers, general reporter molecules, photon traps, UV sinks or
CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
CC fungal species, and in fruits and vegetables to enhance their
CC marketability. CFMs embedded in a gel matrix improve image quality in
CC situations of distorted light spectra (biomatrix). The first all-protein
CC chromophore to be isolated was Green Fluorescent protein (GFP). The
CC sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
CC sequences

XX
SQ Sequence 660 BP; 193 A; 146 C; 161 G; 160 T; 0 U; 0 Other;

Query Match 76.0%; Score 639.2; DB 6; Length 660;
Best Local Similarity 98.0%; Pred. No. 2.5e-142;
Matches 647; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 1 TCCGTTATCGCTTAACAGATGACCTACAAAGTTTATATGTGACGACCGTCAATGACAC 60
Db 1 TCCGTTATCGCTTAACAGATGACCTACAAAGTTTATATGTGACGACCGTCAATGACAC 60
Qy 61 TACTTTGAGGTCGAGGCGATGGAAGAAAGCCTTACGAGGGGAGCAGCGTAAGG 120
Db 61 TACTTTGAGGTCGAGGCGATGGAAGAAAGCCTTACGAGGGGAGCAGCGTAAGG 120
Qy 121 CTGGCTGTCCACAAAGGCGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTGT 180
Db 121 CTCACTGTCCACAAAGGCGGACCTCTGCCATTTGCTTGGGATATTTATCACCACAGTGT 180
Qy 181 CAGTACGGAACATACATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAGCAG 240
Db 181 CAGTACGGAACATACATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAGCAG 240
Qy 241 TCATTCCCGGGGAGATATACATGGAGAGGATCATGAATTTGAAGTGTGCAAGTGTGT 300
Db 241 TCATTCCCTGAGGATATACATGGAGAGGATCATGAATTTGAAGTGTGCAAGTGTGT 300
Qy 301 ACTGTGACATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT 360
Db 301 ACTGTGACATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT 360
Qy 361 GGTTTGAACCTTCCCTCCCAATGGACCTGTTATGCAGAGAGACACAGGCTGGGAACCC 420
Db 361 GGTTTGAACCTTCCCTCCCAATGGACCTGTTATGCAGAGAGACACAGGCTGGGAACCC 420
Qy 421 AACACTGAGCGTCTCTTTGCAGAGATGGAATGCTGATAGGAACAACTTTATGCTCTG 480
Db 421 AACACTGAGCGTCTCTTTGCAGAGATGGAATGCTGATAGGAACAACTTTATGCTCTG 480
Qy 481 AAGTTAGAGAGGCTGCTCACTATTTGTTGTAATTCAAATCTACTTCAAGGCAAGGAAG 540
Db 481 AAGTTAGAGAGGCTGCTCACTATTTGTTGTAATTCAAATCTACTTCAAGGCAAGGAAG 540
Qy 541 CCTGTGAAGATGCCAGGCTATCACTATGTTGACCGCAAACTGGATGAACCAATCACAAC 600
Db 541 CCTGTGAAGATGCCAGGCTATCACTATGTTGACCGCAAACTGGATGAACCAATCACAAC 600
Qy 601 AAGGATTACACTTCCGTTGAGCAGTGTGAAATTTCCATTTGCAACGCAAACTGTGTGCGCC 660
Db 601 AAGGATTACACTTCCGTTGAGCAGTGTGAAATTTCCATTTGCAACGCAAACTGTGTGCGCC 660

RESULT 15

ABZ12159
ID ABZ12159 standard; DNA; 669 BP.
XX
XX
AC ABZ12159;
XX

DT 22-JAN-2003 (first entry)
 DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 217.
 XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunscreen; ds.
 XX Discosoma sp.
 OS W0200270703-A2.
 PN 12-SEP-2002.
 PD 01-MAR-2002; 2002WO-GB000928.
 PF 02-MAR-2001; 2001US-0273227P.
 XX 21-MAR-2001; 2001AU-00003874.
 PR 15-OCT-2001; 2001US-0329816P.
 XX (NUFA-) NUFARM LTD.
 PA (UYOU) UNIV QUEENSLAND.
 PA (JONE/) JONES E L.
 XX Jones EL, Karan M, Brughiera F, Mason J, Dove SG;
 PI Hoegh-Guldberg IO, Prescott M,
 XX WPI; 2002-740765/80.
 DR Novel color-facilitating molecule for producing a biomatrix, has a
 XX polypeptide which alone/along with molecules imparts altered visual
 PT characteristics to cells in the absence of excitation by extraneous non-
 PT white light.
 XX Disclosure; Page 487; 510pp; English.
 PS The invention relates to an isolated colour-facilitating molecule (CFM)
 XX comprising a polypeptide which, in a cell, alone or together with one or
 CC more other molecules imparts an altered visual characteristic to the cell
 CC when visualised by a human eye in the absence of excitation by extraneous
 CC non-white light or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant
 CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development of
 CC new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
 CC sequences
 XX
 QY 121 CTGGCTGTCCAAAGGCGGACCTCTGCCATTTGCTTGGGATATTTTATACCAACAGTGT 180
 DB 124 CTGGCTGTCCAAAGGCGGACCTCTGCCATTTGCTTGGGATATTTTATACCAACAGTGT 183
 QY 181 CAGTACGGAAGCATACCATTCACCAAGTACCCCTGAAGACATCCCTGACTATGTAAAGCAG 240
 DB 184 CAGTACGGAAGCATACCATTCACCAAGTACCCCTGAAGACATCCCTGACTATGTAAAGCAG 243
 QY 241 TCATTCCCGGGGAGATATACATGGGAGAGGATCATGAACCTTTGAAGATGGTGCAGTGTGT 300
 DB 244 TCATTCCCGGGGAGATTTACATGGGAGAGGATCAAGAACTTTGAAGATGGTGCAGTGTGT 303
 QY 301 ACTGTCAAGCAATGATCCAGCATCCAAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT 360
 DB 304 CCTGTGAGCAATGATCCAGCATCCAAAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT 363
 QY 361 GGTTTGAACTTTCTCCCAATGGACCTGTTATGAGAAGAAGACACAGGGCTGGGAACCC 420
 DB 364 GGTTTGAACTTTCTCCCAATGGACCTGTTATGAGAAGAAGACACAGGGCTGGGAACCC 423
 QY 421 AACACTGAGCGTCTCTTTGACACGAGATGGAATGCTGATAGGAAACAACCTTTATGGCTCTG 480
 DB 424 CACTCTGAGCGTCTCTTTGACACGAGATGGAATGCTGATAGGAAACAACCTTTATGGCTCTG 483
 QY 481 AAGTTAGAAGGAGGTGCTCACTATTGTTGTAATTCAAATCTACTTACAAGGCAAGAAG 540
 DB 484 AAGTTAGAAGGAGGCGGTCACTATTGTTGTAATTCAAATCTACTTACAAGGCAAGAAG 543
 QY 541 CCTGTGAAGATGCCAGGGTATCACTATGTTACCGCAAACTGGATTAACCAATCACAAC 600
 DB 544 CCTGTGAAGATGCCAGGGTATCACTATGTTACCGCAAACTGGATTAACCAATCACAAC 603
 QY 601 AAGGATTACACTTCCGTTGACGAGTGTGAAATTTCCATTGCACGCAAACTGTGTGTCGCC 660
 DB 604 AAGGATTACACTTCCGTTGACGAGTGTGAAATTTCCATTGCACGCAAACTGTGTGTCGCC 663

Search completed: August 13, 2004, 17:09:52
 Job time : 410.5 secs

Query Match 76.0%; Score 639.2; DB 6; Length 669;
 Best Local Similarity 98.0%; Pred. No. 2.5e-142;
 Matches 647; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 QY 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATGTCAGGCAACGGTCAATGGACAC 60
 DB 4 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATGTCAGGCAACGGTCAATGGACAC 63
 QY 61 TACTTTGAGGTGAGGGCGATGGAAGAAAGCAAGCCCTTACAGGGGGAGCAGACGGTAAGG 120
 DB 64 TACTTTGAGGTGAGGGCGATGGAAGAAAGCAAGCCCTTACAGGGGGAGCAGACGGTAAGG 123

This Page Blank (uspto)

Db 79 TTTGAAATAGAAAGGCGAAGGAGAGGAGGCCATACGAGGCCACAATACCGTAAAGCTT 138

Db 79 TTTGAAATAGAAAGGCGAAGGAGAGGAGGCCATACGAGGCCACAATACCGTAAAGCTT 138

QY 124 GCTGTCAACAGGGGGAGCTCTGCCATTTGCTTGGGATATTTTATCACCAGTGTGAG 183
 Db |||||
 QY 139 AAGGTAAACCAAGGGGGAGCTTTGCCATTTGCTTGGGATATTTTATCACCAGTGTGAG 198
 Db |||||
 QY 184 TAGGGAAGCATACCAATTACCAAGTACCTCGAAGACATCCCTGAGTATGTAAAGCAGTCA 243
 Db |||||
 QY 199 TATGGAAGCAAGGTATATGTCAAGCATCTCGCCGACATACCAAGCATATAAAAAGCTGTCA 258
 QY |||||
 QY 244 TTCCCGGGGAGATATACATGGGAGAGGATCATGAATTTGAGATGGTGCAGTGTGTACT 303
 Db |||||
 QY 259 TTTCCTGAAGGATTTAAATGGGAAGGTCATGAATTTTGAAGACGGTGGCGTCTGTACT 318
 QY |||||
 QY 304 GTCAACAATGATTCAGCATCCAAGCAACTGTTTTCATCTACCATGTCAAGTTCCTGTGT 363
 Db |||||
 QY 319 GTAAACCCAGGATTTCCAGTTTTCAGGATGCTGTTTCATCTCAAGGTCAAGTTCATTGGC 378
 QY |||||
 QY 364 TTGAATTTCTCCCAATGGAGCTGTATGCAAGAGAGACACAGGGCTGGGAACCCCAAC 423
 Db |||||
 QY 379 GTGAATTTCTCTGTATGGACCTTTATGCAAAAGAGACAAATGGGCTGGGAAGCCAGC 438
 QY |||||
 QY 424 ACTGAGCGTCTCTTTGCAAGATGGAATGCTGATAGGAACAACCTTTATGGCTCTGAAG 483
 Db |||||
 QY 439 ACTGAGCGTCTCTTTGCAAGATGGAATGCTGATAGGAACAACCTTTATGGCTCTGAAG 498
 QY |||||
 QY 484 TTGAAGGAGGTGTCATTTTGTGTAATTTCAAACTATTTTACATGSCAAAGAGCGCT 543
 Db |||||
 QY 499 TTGAAGGAGGTGTCATTTTGTGTAATTTCAAACTATTTTACATGSCAAAGAGCGCT 558
 QY |||||
 QY 544 GTGAGATGCCAGGATATCACTATCTTGCACCGCAAACTGGATGTAACCAATCACAACAG 603
 Db |||||
 QY 559 GTGAGATGCCAGGATATCACTATCTTGCACCGCAAACTGGATGTAACCAATCACAACAG 618
 QY |||||
 QY 604 GATTACACTTCGTTGAGCAGTGTGAAATTTCCATTTGCACGCAAACTGTGTGTC 657
 Db |||||
 QY 619 GACTATACAAATCGTTGAGCAGTGTGAAAGAACCGAGGAGCCACCATCTGTTC 672

RESULT 2

US-10-314-936-3
 ; Sequence 3, Application US/10314936
 ; Publication No. US20040110225A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gibbs, Patrick D.L.
 ; APPLICANT: Carter, Robert W.
 ; APPLICANT: Schmale, Michael C.
 ; TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES
 ; FILE REFERENCE: 638.004
 ; CURRENT APPLICATION NUMBER: US/10/314,936
 ; CURRENT FILING DATE: 2002-12-09
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 711
 ; TYPE: DNA
 ; ORGANISM: Artificial
 ; FEATURE:
 ; OTHER INFORMATION: mutant red fluorescent protein
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(711)
 ; OTHER INFORMATION:

Query Match 46.2%; Score 388.4; DB 17; Length 711;
 Best Local Similarity 74.6%; Pred. No. 1e-93;
 Matches 488; Conservative 0; Mismatches 166; Indels 0; Gaps 0;
 QY 4 GTTATCGCTAAACAGATGACCTTACAAAGTTTATGTGTCAGGACCGTCAATGGACACTAC 63
 Db |||||
 QY 19 GTTATCAAGGAGTTTCATGAGGTTTAAAGTTTGTGATGGAAGGAACCGTCAATGGGCACGAG 78
 Db |||||
 QY 64 TTTGAGTTCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGGAGCAGACCGTAAAGCTG 123

Db |||||
 QY 79 TTTGAATAGAGGCGAAGGAGGAGCCATACGAGGCGCAATACCGTAAAGCTT 138
 QY |||||
 QY 124 GCTGTCAACAGGGGGAGCTCTGCCATTTGCTTGGGATATTTTATCACCAGTGTGAG 183
 Db |||||
 QY 139 AAGGTAAACCAAGGGGGAGCTTTGCCATTTGCTTGGGATATTTTGTCAACAATTTGAG 198
 QY |||||
 QY 184 TAGGGAAGCATACCAATTACCAAGTACCTCGAAGACATCCCTGAGTATGTAAAGCAGTCA 243
 Db |||||
 QY 199 TATGGAAGCAAGGTATATGTCAAGCATCTCGCCGACATACCAAGCATATAAAAAGCTGTCA 258
 QY |||||
 QY 244 TTCCCGGGGAGATATACATGGGAGAGGATCATGAATTTTGAAGATGGTGCAGTGTGTACT 303
 Db |||||
 QY 259 TTTCCTGAAGGATTTAAATGGGAAGGTCATGAATTTTGAAGACGGTGGCGTCTGTACT 318
 QY |||||
 QY 304 GTCAACAATGATTCAGCATCCAAGCAACTGTTTTCATCTACCATGTCAAGTTCCTGTGT 363
 Db |||||
 QY 319 GTAAACCCAGGATTTCCAGTTTTCAGGATGCTGTTTTCATCTCAAGGTCAAGTTCATTGGC 378
 QY |||||
 QY 364 TTGAATTTCTCCCAATGGAGCTGTATGCAAGAGAGACACAGGGCTGGGAACCCCAAC 423
 Db |||||
 QY 379 GTGAATTTCTCTGTATGGACCTTTATGCAAAAGAGACAAATGGGCTGGGAAGCCAGC 438
 QY |||||
 QY 424 ACTGAGCGTCTCTTTGCAAGATGGAATGCTGATAGGAACAACCTTTATGGCTCTGAAG 483
 Db |||||
 QY 439 ACTGAGCGTCTCTTTGCAAGATGGAATGCTGATAGGAACAACCTTTATGGCTCTGAAG 498
 QY |||||
 QY 484 TTGAAGGAGGTGTCATTTTGTGTAATTTCAAACTATTTTACATGSCAAAGAGCGCT 543
 Db |||||
 QY 499 TTGAAGGAGGTGTCATTTTGTGTAATTTCAAACTATTTTACATGSCAAAGAGCGCT 558
 QY |||||
 QY 544 GTGAGATGCCAGGATATCACTATCTTGCACCGCAAACTGGATGTAACCAATCACAACAG 603
 Db |||||
 QY 559 GTGAGATGCCAGGATATCACTATCTTGCACCGCAAACTGGATGTAACCAATCACAACAG 618
 QY |||||
 QY 604 GATTACACTTCGTTGAGCAGTGTGAAATTTCCATTTGCACGCAAACTGTGTGTC 657
 Db |||||
 QY 619 GACTATACAAATCGTTGAGCAGTGTGAAAGAACCGAGGAGCCACCATCTGTTC 672

RESULT 3

US-10-332-733-22
 ; Sequence 22, Application US/10332733
 ; Publication No. US20040106565A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Margarete Odenthal and Diana Jung
 ; TITLE OF INVENTION: Gene Expression, Genome Alteration And Reporter Expression
 ; FILE REFERENCE: 1472/68806
 ; CURRENT APPLICATION NUMBER: US/10/332,733
 ; CURRENT FILING DATE: 2003-06-10
 ; NUMBER OF SEQ ID NOS: 41
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 22
 ; LENGTH: 666
 ; TYPE: DNA
 ; ORGANISM: Artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of artificial sequence: Red
 ; OTHER INFORMATION: Fluorescent Protein

Query Match 45.6%; Score 383.6; DB 17; Length 666;
 Best Local Similarity 74.2%; Pred. No. 1.9e-92;
 Matches 485; Conservative 0; Mismatches 169; Indels 0; Gaps 0;
 QY 4 GTTATCGCTAAACAGATGACCTTACAAAGTTTATGTGTCAGGACCGTCAATGGACACTAC 63
 Db |||||
 QY 7 GTTATCAAGGAGTTTCATGAGGTTTAAAGTTTCGCATGGAAGGAACCGTCAATGGGCACGAG 66
 Db |||||
 QY 64 TTTGAGTTCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGGAGCAGACCGTAAAGCTG 123
 Db |||||
 QY 67 TTTGAATAGAAAGCGGAAGGAGGAGGGAGGCCATACGAAGGCCCAATACCGTAAAGCTT 126

QY 124 GCTGTCACCAAGGGCGGACCTCTGCCATTTCGTTGGGATATTTTATCACCACAGTGTCTAG 183
DB 127 AAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCCACCAAAATTCAG 186
QY 184 TACGGAAGCATACCAATTCACCAAGTACCTCGAAGACATCCCTCGACTATGTATAAGCAGTCA 243
DB 187 TATGGAAGCAAGGTATATGTCAAGCACCTCGCGACATACCAGACTATAAAAAGCTGTCA 246
QY 244 TTCCCGGGAGATATACATGGGAGAGGATCATGAACCTTTGAAGATGGTGCAGTGTGTACT 303
DB 247 TTTCCTGAAGGATTTAAATGGGAAAGGGTTCATGAACCTTTGAAGACGGTGGCGTCTGTACT 306
QY 304 GTCCAGCATGATTCAGCATCAAGGCACTGTTTCATCTACCATGTCAAGTTCCTCTGCT 363
DB 307 GTAACCCAGGATTCAGTTTGCAGATGGCTGTTTCATCTCAAGGTCAAGTTCATGTGC 366
QY 364 TTGAACCTTTCTCCCAATGGACCTGTTATGCAGAAAGACACACAGGGCTGGAAACCCAAAC 423
DB 367 GTGAACCTTTCTCCGATGGACCTGTTATGCAGAAAGACACAAATGGCTGGGAAGCCAGC 426
QY 424 ACTGAGCGTCTTTGACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGGAAG 483
DB 427 ACTGAGCGTCTTTATCCTCGTATGGCGGTGTTGAAAGGAGAGATTTCATAAGGCTCTGGAAG 486
QY 484 TTGAAGGAGGTGGTCACTATTGTTGTGAATTCAAATCTACTTACAAGGCAAGAGAGGCT 543
DB 487 CTGAAGAGCGTGGTCACTTACCTAGTTGAAATTCAAAAGTATTTATGGAAGAGAGGCT 546
QY 544 GTGAAGATGCCAGGATACATGATGTTGACCGCAAACTGGATGTAAACCAATCACACAAAG 603
DB 547 GTGCAGCTACCAAGGTACTACTATGTTGACTCCAACTGATATACCAAGCCACACAAAGAA 606
QY 604 GATTACACTTCGTTGAGCAGTGTGAATTTCCATTGACGCAAACTGTGTGTC 657
DB 607 GACTATACAATCGTTGAGCAGTATGAAGAAGACCGAGGAGCGCCACCATCTGTTC 660

RESULT 4

US-09-967-772-6
; Sequence 6, Application US/09967772
; Patent No. US20020164577A1
; GENERAL INFORMATION:
; APPLICANT: TSIEN, Roger
; APPLICANT: GONZALEZ, Jesus
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS
; FILE REFERENCE: REGEN1290-5
; CURRENT APPLICATION NUMBER: US/09/967,772
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 09/459,956
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 08/765,860
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: PCT/US96/09652
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: US 08/481,977
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; TYPE: DNA
; ORGANISM: Discosoma sp "red"
US-09-967-772-6
Query Match 45.6%; Score 383.6; DB 9; Length 678;
Best Local Similarity 74.2%; Pred. No. 2e-92;
Matches 485; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

QY 64 TTTGAGTCTGAGGCGATGAAAGGAAAGCCTTACGAGGGGAGCAGACGCTAAGGCTG 123
DB 79 TTTGAAATAGAAAGGCGAGAGGGAGGCCATACGAGGGCCAAATACCGTTAAGCTT 138
QY 124 GCTGTCACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGTCTAG 183
DB 139 AAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACCAAAATTCAG 198
QY 184 TACGGAAGCATACCAATTCACCAAGTACCTCGAAGACATCCCTCGACTATGTATAAGCAGTCA 243
DB 199 TATGGAAGCAAGGTATATGTCAAGCACCTCGCGACATACCAGACTATAAAAAGCTGTCA 258
QY 244 TTCCCGGGAGATATACATGGGAGAGGATCATGAACCTTTGAAGATGGTGCAGTGTGTACT 303
DB 259 TTTCCTGAAGGATTTAAATGGGAAAGGGTTCATGAACCTTTGAAGACGGTGGCGTCTGTACT 318
QY 304 GTCCAGCATGATTCAGCATCCAGGCAACTGTTTTCATCTACCATGTCAAAGTTCCTCTGCT 363
DB 319 GTAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTCAAGGTCAAGTTCATTGGC 378
QY 364 TTGAACCTTTCTCCCAATGGACCTGTTATGCAGAAAGACACACAGGGCTGGAAACCCAAAC 423
DB 379 GTGAACCTTTCTCCGATGGACCTGTTATGCAGAAAGACAAATGGCTGGGAAGCCAGC 438
QY 424 ACTGAGCGTCTTTGACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGGAAG 483
DB 439 ACTGAGCGTCTTTGATCTCGTATGGCGGTGTTGAAAGGAGAGATTTCATAAGGCTCTGGAAG 498
QY 484 TTGAAGGAGGTGGTCACTATTGTTGTGAATTCAAATCTACTTACAAGGCAAGAGAGGCT 543
DB 499 CTGAAGAGCGTGGTCACTTACCTAGTTGAATTCAAAAGTATTTACATGGCAAGAGGCT 558
QY 544 GTGAAGATGCCAGGATACATGATGTTGACCGCAAACTGGATGTAAACCAATCACACAAAG 603
DB 559 GTGCAGCTACCAAGGTACTACTATGTTGACTCCAACTGGATATACCAAGCCACCAACGAA 618
QY 604 GATTACACTTCGTTGAGCAGTGTGAATTTCCATTGACGCAAACTGTGTGTC 657
DB 619 GACTATACAATCGTTGAGCAGTATGAAGAAGACCGAGGAGCGCCACCATCTGTTC 672

RESULT 5

US-10-132-067-3
; Sequence 3, Application US/10132067
; Publication No. US2003020355A1
; GENERAL INFORMATION:
; APPLICANT: Bradbury, Andrew
; APPLICANT: Zeytun, Ahmet
; APPLICANT: Waldo, Geoffrey
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Fluorobodies: Binding Ligands With Intrinsic
; TITLE OF INVENTION: Fluorescence
; FILE REFERENCE: 021362-000600US
; CURRENT APPLICATION NUMBER: US/10/132,067
; CURRENT FILING DATE: 2002-04-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; OTHER INFORMATION: red fluorescent protein (dsRED)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(678)
; OTHER INFORMATION: dsRED
US-10-132-067-3

Query Match 45.6%; Score 383.6; DB 13; Length 678;
Best Local Similarity 74.2%; Pred. No. 2e-92;
Matches 485; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

```
QY 4 GTTATCGCTAAACAGATGACCTACAAAGTTTATATGTTCAGGCACCGTCAATGACACTAC 63
DB 19 GTTATCAAGGAGTTTCATGAGTTTAAAGTTCGATGGAAGAACCGTCAATGGGCACGAG 78
QY 64 TTTGAGGTGCAAGGCGATGGAAGAAAGCCCTTACGAGGGGGAGACGACCGTAAAGCTG 123
DB 79 TTTGAAATAGAGGCGAAGGAGGAGGGGAGGCCATACGAAGGCCCAATACCGTAAAGCTT 138
QY 124 GCTGTCACCAAGGGGGACCTCTGCCATTTGCTTGGGATATTTTATCACCAGTGTCTG 183
DB 139 AAGGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTACCACAATTTGAG 198
QY 184 TACGGAAGCATACCAATTCACCAAGTACCTTGAAGACATCCCTGACTATGTAAAGCAGTCA 243
DB 199 TATGGAACCAAGGTATATGTCAAGCACCTTCCGACATACCAAGACTATAAAGGCTGCA 258
QY 244 TTTCCGGGGAGATATACATGGGAGGAGTATGTAAGTGTGAGTGTGAGTGTGAGTGTG 303
DB 259 TTTCCCTGAAGGATTTAAATGGGAAAGGGTTCATGAACTTTGAAGACGGTGGCGTCTG 318
QY 304 GTCGCAATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAAAGTTCCTCTGT 363
DB 319 GTAACCCAGGATTCAGTTCGAGATGGCTGTTTCATCTCAAGGTCAAGTTCATTTGGC 378
QY 364 TTGAACCTTCTCCCAATGGACCTGTTATGAGAAAGACACAGGGGCTGGGAACCCAAAC 423
DB 379 GTGAACCTTCTCCGATGGACCTGTTATGCAAAAGAAAGACAATGGGCTGGGAAGCCAGC 438
QY 424 ACTGAGCGTCTTTGACGAGATGGAATGCTGATAGGAAACACTTTATGGCTCTGAG 483
DB 439 ACTGAGCGTCTTTGATCCTCGTGGTGGCTGTTGAAAGGAGAGATTCATAAGGCTCTG 498
QY 484 TTGAAGGAGGTGGTCACTATTGTTGTAATTCAAATCTACTTCAAGGCAAAAGAGCCT 543
DB 499 CTGAAGACGGTGGTCACTTACCTAGTTGAAATTCAAAGTATTTACATGCAAGAGCCT 558
QY 544 GTGAAGATGCCAGGATATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAAACAG 603
DB 559 GTGACGCTACCAAGGATCACTATGTTGACTCCAAACTGGATATACCAAGCCACCAACGAA 618
QY 604 GATTACACTTCGTTGAGCAGTGAATTTCCATTTGACGCAAACTGTGGTC 657
DB 619 GACTATACATCGTTGAGCAGTATGAAAGAACCGGAGGACGCCACCATCTGTTC 672
```

RESULT 6

US-10-006-922-11

```
; Sequence 11, Application US/10006922
; Publication No. US20020197676A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey A
; APPLICANT: Fradkov, Arcady F.
; APPLICANT: Labas, Yulii A.
; APPLICANT: Matz, Mikhail V.
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and
; FILE REFERENCE: CLON-035CIP
; CURRENT APPLICATION NUMBER: US/10/006,922
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/120,330
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/457,898
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,144
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,477
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/457,556
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/444,338
; PRIOR FILING DATE: 1999-11-19
```

```
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma species
US-10-006-922-11
```

```
Query Match 45.6%; Score 383.6; DB 14; Length 678;
Best Local Similarity 74.2%; Pred. No. 2e-92; Mismatches 169; Indels 0; Gaps 0;
Matches 485; Conservative 0;
```

```
QY 4 GTTATCGCTAAACAGATGACCTACAAAGTTTATATGTTCAGGCACCGTCAATGACACTAC 63
DB 19 GTTATCAAGGAGTTTCATGAGTTTAAAGTTCGATGGAAGAACCGTCAATGGGCACGAG 78
QY 64 TTTGAGGTGCAAGGCGATGGAAGAAAGCCCTTACGAGGGGGAGACGACCGTAAAGCTG 123
DB 79 TTTGAAATAGAGGCGAAGGAGGAGGGGAGGCCATACGAAGGCCCAATACCGTAAAGCTT 138
QY 124 GCTGTCACCAAGGGGGACCTCTGCCATTTGCTTGGGATATTTTATCACCAGTGTCTG 183
DB 139 AAGGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTACCACAATTTGAG 198
QY 184 TACGGAAGCATACCAATTCACCAAGTACCTTGAAGACATCCCTGACTATGTAAAGCAGTCA 243
DB 199 TATGGAACCAAGGTATATGTCAAGCACCTTCCGACATACCAAGACTATAAAGGCTGCA 258
QY 244 TTTCCGGGGAGATATACATGGGAGGAGTATGTAAGTGTGAGTGTGAGTGTGAGTGTG 303
DB 259 TTTCCCTGAAGGATTTAAATGGGAAAGGGTTCATGAACTTTGAAGACGGTGGCGTCTG 318
QY 304 GTCGCAATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAAAGTTCCTCTGT 363
DB 319 GTAACCCAGGATTCAGTTCGAGATGGCTGTTTCATCTCAAGGTCAAGTTCATTTGGC 378
QY 364 TTGAACCTTCTCCCAATGGACCTGTTATGAGAAAGACACAGGGGCTGGGAACCCAAAC 423
DB 379 GTGAACCTTCTCCGATGGACCTGTTATGCAAAAGAAAGACAATGGGCTGGGAAGCCAGC 438
QY 424 ACTGAGCGTCTTTGACGAGATGGAATGCTGATAGGAAACACTTTATGGCTCTGAG 483
DB 439 ACTGAGCGTCTTTGATCCTCGTGGTGGCTGTTGAAAGGAGAGATTCATAAGGCTCTG 498
QY 484 TTGAAGGAGGTGGTCACTATTGTTGTAATTCAAATCTACTTCAAGGCAAAAGAGCCT 543
DB 499 CTGAAGACGGTGGTCACTTACCTAGTTGAAATTCAAAGTATTTACATGCAAGAGCCT 558
QY 544 GTGAAGATGCCAGGATATCACTATGTTGACCGCAAACTGGATGTAACCAATCACAAACAG 603
DB 559 GTGACGCTACCAAGGATCACTATGTTGACTCCAAACTGGATATACCAAGCCACCAACGAA 618
QY 604 GATTACACTTCGTTGAGCAGTGAATTTCCATTTGACGCAAACTGTGGTC 657
DB 619 GACTATACATCGTTGAGCAGTATGAAAGAACCGGAGGACGCCACCATCTGTTC 672
```

RESULT 7

US-10-121-258-2

```
; Sequence 2, Application US/10121258
; Publication No. US20030059835A1
; GENERAL INFORMATION:
; APPLICANT: Campbell, Robert
; APPLICANT: Tsien, Roger
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; FILE REFERENCE: UC083.1CF2CPI
; CURRENT APPLICATION NUMBER: US/10/121,258
; PRIOR FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
```

```
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(678)
; OTHER INFORMATION: wild-type DeRed
US-10-121-258-2

Query Match      45.6%; Score 383.6; DB 15; Length 678;
Best Local Similarity 74.2%; Pred. No. 2e-92;
Matches 485; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

QY 4 GTTATCGCTAAACAGATGACCTACAAAGTTTATATGTTCAGGACCGTCAATGGACACTAC 63
Db 19 GTTATCAAGGAGTTTCATGAGGTTTAAAGTTTCGCATGGAAGAAACCGTCAATGGCACGAG 78
QY 64 TTTGAGGTCGAAGGCGATGGAAGAAAGAAAGCCTTACGAGGGGGGAGCAGACCGTAAAGGCTG 123
Db 79 TTTGAAATAGAGGCGAAGGAGAGGGAGGCGCATACGAAGGCCCAATACCGTAAAGCTT 138
QY 124 GCTGTACCAAGGGCGGACCTCTGCCATTGCTTGGGATATTTTATCACCACAGTGTGCTG 183
Db 139 AAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATTTTGTACCAACAATTTTCA 198
QY 184 TACGGAACATACCAATTCACCAAGTACCTGAGACATCCCTGACTGCTATGTAAAGCAGTCA 243
Db 199 TATGGAAGCAAGGTATATGTCAAGCACCTTCCGACATACCAAGCTATAAAAAGCTGTCA 258
QY 244 TTCCCGGGAGATATACATGCGAGAGGATCATGAACCTTTGAAGATGCTGAGTGTGCTACT 303
Db 259 TTTCTGGAAGGATTTAAATGGGAAGGGTCATGAACCTTTGAAGACGGTGGCGTCTGTACT 318
QY 304 GTCAGCAATGATTCAGCATCCAAAGCAACTGTTTTCATCTACCATGTCAAGTTCTCTGCT 363
Db 319 GTAACCCAGGATTCAGTTTCAGAGATGGCTGTTTCATCTACAAGTCAAGTTCATTGGC 378
QY 364 TTGAACCTTCTCCCAATGGACCTGTTATGAGAAGAGACACAGGGCTGGGAACCCCAAC 423
Db 379 GTGAACCTTCTCCATGGACCTGTTATGCAAAAGAGACAAATGGGCTGGGAAGCCAGC 438
QY 424 ACTGAGCCTCTTTGACAGAGATGGAATGCTGATAGGAACAACCTTTATGGCTCTGAAG 483
Db 439 ACTGAGCCTTTGATPCTCGTGGCTGTTGAAAGGAGAGATTCTAAGGCTCTGAAG 498
QY 484 TTAGAAGAGAGTGGTCACTATTGTTGTGAATCTAAATCTACTTACAGGCAAAAGAGCCT 543
Db 499 CTGAAAGACGGTGGTCAATTACCTAGTTGAATTTCAAAAGTATTTACATGGCAAGAGCCT 558
QY 544 GTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAAACCAATCAACAAG 603
Db 559 GTGACGCTACCAAGGCTACTATGTTGACTCCAACTGGATATTAACAAGCCACCAACGAA 618
QY 604 GATTACACTTCGTTGAGCAGTGTCAAAATTTCCATTGTCACGCAAAACCTGTGCTC 657
Db 619 GACTATACATCGTTGAGCAGTATGAAGAACCGAGGGACGCCACCATCTGTTC 672

RESULT 8
US-10-335-517-6
; Sequence 6, Application US/10335517
; Publication No. US20030207248A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSJEN, Roger
; APPLICANT: GONZALEZ, Jesus
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS
; FILE REFERENCE: REGEN1290-5
; CURRENT APPLICATION NUMBER: US/10/335,517
; CURRENT FILING DATE: 2002-12-31
```

```
; PRIOR APPLICATION NUMBER: US/09/967,772
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 09/459,956
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 08/765,860
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: PCT/ US96/09652
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: US 08/481,977
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp "red"
US-10-335-517-6

Query Match      45.6%; Score 383.6; DB 16; Length 678;
Best Local Similarity 74.2%; Pred. No. 2e-92;
Matches 485; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

QY 4 GTTATCGCTAAACAGATGACCTACAAAGTTTATATGTTCAGGACCGTCAATGGACACTAC 63
Db 19 GTTATCAAGGAGTTTCATGAGGTTTAAAGTTTCGCATGGAAGAAACCGTCAATGGCACGAG 78
QY 64 TTTGAGGTCGAAGGCGATGGAAGAAAGAAAGCCTTACGAGGGGGGAGCAGACCGTAAAGGCTG 123
Db 79 TTTGAAATAGAGGCGAAGGAGAGGGAGGCGCATACGAAGGCCCAATACCGTAAAGCTT 138
QY 124 GCTGTACCAAGGGCGGACCTCTGCCATTGCTTGGGATATTTTATCACCACAGTGTGCTG 183
Db 139 AAGGTAACCAAGGGGGACCTTTGCCATTGCTTGGGATATTTTGTACCAACAATTTTCA 198
QY 184 TACGGAACATACCAATTCACCAAGTACCTGAGACATCCCTGACTGCTATGTAAAGCAGTCA 243
Db 199 TATGGAAGCAAGGTATATGTCAAGCACCTTCCGACATACCAAGCTATAAAAAGCTGTCA 258
QY 244 TTCCCGGGAGATATACATGCGAGAGGATCATGAACCTTTGAAGATGCTGAGTGTGCTACT 303
Db 259 TTTCTGGAAGGATTTAAATGGGAAGGGTCATGAACCTTTGAAGACGGTGGCGTCTGTACT 318
QY 304 GTCAGCAATGATTCAGCATCCAAAGCAACTGTTTTCATCTACCATGTCAAGTTCTCTGCT 363
Db 319 GTAACCCAGGATTCAGTTTCAGAGATGGCTGTTTCATCTACAAGTCAAGTTCATTGGC 378
QY 364 TTGAACCTTCTCCCAATGGACCTGTTATGAGAAGAGACACAGGGCTGGGAACCCCAAC 423
Db 379 GTGAACCTTCTCCATGGACCTGTTATGCAAAAGAGACAAATGGGCTGGGAAGCCAGC 438
QY 424 ACTGAGCCTCTTTGACAGAGATGGAATGCTGATAGGAACAACCTTTATGGCTCTGAAG 483
Db 439 ACTGAGCCTTTGATPCTCGTGGCTGTTGAAAGGAGAGATTCTAAGGCTCTGAAG 498
QY 484 TTAGAAGAGAGTGGTCACTATTGTTGTGAATCTAAATCTACTTACAGGCAAAAGAGCCT 543
Db 499 CTGAAAGACGGTGGTCAATTACCTAGTTGAATTTCAAAAGTATTTACATGGCAAGAGCCT 558
QY 544 GTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAAACCAATCAACAAG 603
Db 559 GTGACGCTACCAAGGCTACTATGTTGACTCCAACTGGATATTAACAAGCCACCAACGAA 618
QY 604 GATTACACTTCGTTGAGCAGTGTCAAAATTTCCATTGTCACGCAAAACCTGTGCTC 657
Db 619 GACTATACATCGTTGAGCAGTATGAAGAACCGAGGGACGCCACCATCTGTTC 672

RESULT 9
US-10-334-288-6
; Sequence 6, Application US/10334288
; Publication No. US20040002123A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
```

APPLICANT: TSJEN, Roger
APPLICANT: GONZALEZ, Jesus
TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS
FILE REFERENCE: REGEN1290-5
CURRENT APPLICATION NUMBER: US/10/334,288
CURRENT FILING DATE: 2002-12-31
PRIOR APPLICATION NUMBER: US/09/967,772
PRIOR FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: US 09/459,956
PRIOR FILING DATE: 1999-12-13
PRIOR APPLICATION NUMBER: US 08/765,860
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: PCT/US96/09652
PRIOR FILING DATE: 1996-06-06
PRIOR APPLICATION NUMBER: US 08/481,977
PRIOR FILING DATE: 1995-06-07
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 678
TYPE: DNA
ORGANISM: Discosoma sp "red"
US-10-334-288-6

Query Match 45.6%; Score 383.6; DB 16; Length 678;
Best Local Similarity 74.2%; Pred. No. 2e-92;
Matches 485; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

```
Qy 4 GTTATCGCTTAACAGATGACCTACAAAGTTTATATGTCAGGCGGTCATGGACACTAC 63
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 19 GTTATCAAGGAGTTTCATGAGGTTTAAAGTTTCGCATGGAAGGACCGTCAATGGGCACGAG 78
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 64 TTTGAGGTTCGAAGCGCATGGAAGGAAAGCCTTACGAGGGGGAGCAGCGTGAAGGCTG 123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 79 TTTGAATAGAGCGGAGAGAGGGGCGCCATACGAGGCCACATACCGTAAAGCTT 138
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 124 GCTGTACCAAGGGGACCTCTGCCATTTGCTTGGGATTTTATACCAACAGTGTGAG 183
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 139 AAGGTAACCAAGGGGGACCTTTCGCAATTTGCTTGGGATATTTTGTCAACCAATTTTCAG 198
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 184 TACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGACTATGTTAAAGCAGTCA 243
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 199 TATGGAAGCAAGGTATATGTCAAGACCTCCGACATACCAAGACTATAAAAGCTGTCA 258
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 244 TTCCCGGGGAGATATACATGGGAGGAGTCAATGAACCTTTGAAGATGTTGCAAGTGTGACT 303
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 259 TTTCTCGAAGGATTTAAATGGGAAAGGTCATGAACCTTGAAGCGTGGCGTCTGTTACT 318
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 304 GTCAGCAATGATTCAGCATCCAAAGGCAACTGTTTTCATCTACCATGTCAAGTTCCTGGGT 363
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 319 GTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTTCATCTACAAGGTCAAGTTCATTGGC 378
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 364 TTGAACCTTCTCCCAATGGACCTGTTATGAGAGAGACACAGGCGTGGGAACCCAAAC 423
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 379 GTGAACCTTCTCCGATGGACCTGTTATGCAAAAGAGACATGGGCTGGGAAGCCAGC 438
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 424 ACTGAGCGTCTCTTTGCAAGAGATGGAATGCTGTAGGAAACAACCTTTATGGCTCTGAAAG 483
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 439 ACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAGATTCATAAGGCTCTGAAAG 498
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 484 TTAGAAGGAGTGTGCTACTATTTGTGGAATTCAAATCTACTTACAGGGCAAGAGGCT 543
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 499 CTGAAAGACGGTGTGCTACTTACCTAGTTGAATTCAAAAGTATTTTACATGGCAAGAGCCT 558
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 544 GTGAAGATGCCAGGATATCACTATGTTGACCGCAAACTGGATGTAAACCAATCAACAAG 603
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 559 GTGAGCTACAGGGTACTACTATGTTGACTTCCAACTCGATATAAAGCCACCAAGCAA 618
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 604 GATTACACTTCCGTTGAGCAGTGTGAATTTCCATTTGCAACGCAAACTGTGGTC 657
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 619 GACTATACATCGTTGAGCAGTATGAAGAACCGGAGGACGCCACCATCTGTTTC 672
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 10
US-09-999-745-66
Sequence 66, Application US/09999745
Patent No. US20020157120A1
GENERAL INFORMATION:
APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
APPLICANT: Tsien, Roger Y.
APPLICANT: Baird, Geoffrey
TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS
FILE REFERENCE: REGEN1470-1
CURRENT APPLICATION NUMBER: US/09/999,745
CURRENT FILING DATE: 2001-10-23
PRIOR APPLICATION NUMBER: 09/316,920
PRIOR FILING DATE: 1999-05-21
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn version 3.0
SEQ ID NO 66
LENGTH: 859
TYPE: DNA
ORGANISM: Discosoma sp.
FEATURE:
NAME/KEY: CDS
LOCATION: (54)..(731)
US-09-999-745-66

Query Match 45.6%; Score 383.6; DB 9; Length 859;
Best Local Similarity 74.2%; Pred. No. 2.2e-92;
Matches 485; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

```
Qy 4 GTTATCGCTTAACAGATGACCTACAAAGTTTATATGTCAGGCGGTCATGGACACTAC 63
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 72 GTTATCAAGGAGTTTCATGAGGTTTAAAGTTTCGCATGGAAGGACCGTCAATGGGCACGAG 131
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 64 TTTGAGGTTCGAAGCGCATGGAAGGAAAGCCTTACGAGGGGGAGCAGCGTGAAGGCTG 123
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 132 TTTGAATAGAGCGGAGAGAGGGGAGGCCATACGAAGGCCAATATACCGTAAAGCTT 191
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 124 GCTGTACCAAGGGGACCTCTGCCATTTGCTTGGGATTTTATCACCACAGTGTGAG 183
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 192 AAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACCAATTTTCAG 251
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 184 TACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGACTATGTTAAAGCAGTCA 243
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 252 TATGGAAGCAAGGTATATGTCAAGACCTCCGACATACCAAGACTATAAAAGCTGTCA 311
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 244 TTCCCGGGGAGATATACATGGGAGGAGTCAATGAACCTTTGAAGATGTTGCAAGTGTGACT 303
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 312 TTTCTCGAAGGATTTAAATGGGAAAGGTCATGAACCTTGAAGCGTGGCGTCTGTTACT 371
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 304 GTCAGCAATGATTCAGCATCCAAAGGCAACCTTTTCATCTACCATGTCAAGTTCCTGGGT 363
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 372 GTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAGTTCATTGGC 431
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 364 TTGAACCTTCTCCCAATGGACCTGTTATGAGAGAGACACAGGCGTGGGAACCCAAAC 423
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 432 GTGAACCTTCTCCGATGGACCTGTTATGCAAAAGAGACAAATGGGCTGGGAAGCCAGC 491
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 424 ACTGAGCGTCTCTTTGCAAGAGATGGAATGCTGTAGGAAACAACCTTTATGGCTCTGAAAG 483
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 492 ACTGAGCGTTTGTATCCTCGTGATGGCGTGTGAAAGGAGAGATTCATAAGGCTCTGAAAG 551
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 484 TTAGAAGGAGTGTGCTACTATTTGTGGAATTCAAATCTACTTACAGGGCAAGAGGCT 543
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 552 CTGAAAGACGGTGTGCTACTTACCTAGTTGAATTCAAAAGTATTTTACATGGCAAGAGCCT 611
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 544 GTGAAGATGCCAGGATATCACTATGTTGACCGCAAACTGGATGTAAACCAATCAACAAG 603
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 612 GTGAGCTACAGGGTACTACTATGTTGACTTCCAACTGGATATAAAGCCACCAAGCAA 671
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 604 GATTACACTTCCGTTGAGCAGTGTGAATTTCCATTTGCAACGCAAACTGTGGTC 657
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 672 GACTATACATCGTTGAGCAGTATGAAGAACCGGAGGACGCCACCATCTGTTTC 725
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 11
 US-09-866-538-11
 ; Sequence 11, Application US/09866538
 ; Publication No. US20030032088A1
 ; GENERAL INFORMATION:
 ; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
 ; APPLICANT: TSUEN, Roger
 ; APPLICANT: Campbell, Robert
 ; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
 ; FILE REFERENCE: REGEN1530-2
 ; CURRENT APPLICATION NUMBER: US/09/866,538
 ; CURRENT FILING DATE: 2001-05-24
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 11
 ; LENGTH: 859
 ; TYPE: DNA
 ; ORGANISM: *Drosophila* sp.
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (54)..(731)
 US-09-866-538-11
 Query Match 45.6%; Score 383.6; DB 10; Length 859;
 Best Local Similarity 74.2%; Pred. No. 2.2e-92;
 Matches 485; Conservative 0; Mismatches 169; Indels 0; Gaps 0;
 QY 4 GTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGTCAGGACGGTCAATGGACACTAC 63
 Db 72 GTTATCAAGGAGTTTCATGAGGTTTAAAGTTTCGCATGGAAGAACGGTCAATGGGCACGAG 131
 QY 64 TTTGAGGTCGAAGGCGATGGAAGAAAGAAAGCTTACGAGGGGAGCAGACGGTAAAGCTG 123
 Db 132 TTTGAAATAGAGGCGAAGGAGAGGGAGGCCATACGAGGCCCAATACCGTAAAGCTT 191
 QY 124 GCTGTCAACCAAGGGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACACAGTGTGAG 183
 Db 192 AAGGTAACCAAGGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACCAATTTTCAG 251
 QY 184 TAGCGAAGCATACCATTCACCAAGTACCCTGNAGACATCCCTGACTATGTAAAGCAGTCA 243
 Db 252 TATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATATAAAGCTGTCA 311
 QY 244 TTCCCGGGGAGATATACATGCGAGAGGATCATGAACCTTTGAAGATGGTGCAAGTGTGACT 303
 Db 312 TTTCTGAGGATTTAAATGGGAAGGGTCAATGAATTTGAAGCGGTGGCGTCTGTTACT 371
 QY 304 GTCAGCAATGATTCAGCATCCAGGCAACTGTTTCACTACCATGTCAAGTTCCTGTGGT 363
 Db 372 GTAAACCCAGGATTCAGATTGCGAGATGGCTGTTTTCATCTACAAGGTCAAGTTCATTGGC 431
 QY 364 TTGAACCTTCTCCCAATGGACCTGTATGCAAGAAGAGACACAGGGCTGGGAACCCCAAC 423
 Db 432 GTGAACCTTCTCCGATGGAACCTGTATGCAAGAAGAGACATGGGCTGGGAGCCAGC 491
 QY 424 ACTGAGCGTCTCTTTGCAAGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAG 483
 Db 492 ACTGAGCGTTTGTATCTCTCGTATGGCGTGTGAAAGGAGAGATTCATAAGGCTCTGAAG 551
 QY 484 TTAGAGAGGAGTGGTCACTATTTGTGTAATTCATCTACTTACAGGCGAAGAGGCT 543
 Db 552 CTGAAGAGCGGTGGTCACTTACCTAGTTGTAATTCAAAAGTATTTTACATGGCAAGAGGCT 611
 QY 544 GTGAAGATGCCAGGATATCACTATGTTGACCGCAAACTGGATGTAAACCAATCAACAAG 603
 Db 612 GTGACGCTACCAAGGTTACTATGTTGACTCCAACTGGATATTAACAAGCCACACGAA 671
 QY 604 GATTACACTTCGTTGAGCAGTGTGAATTTCCATTTGCAACGCAACCTGTGGTC 657
 Db 672 GACTATACAACTCGTTGAGCAGTATGAAGAAGCCGAGGAGCGCCACCTCTGTTTC 725

RESULT 13

US-09-865-291-11
; Sequence 11, Application US/09865291
; Publication No. US20030186229A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSUEN, Roger
; APPLICANT: TING, Alice
; APPLICANT: ZHANG, Jin
; TITLE OF INVENTION: EMISSION RATIO-METRIC INDICATORS OF PHOSPHORYLATION
; FILE REFERENCE: REGEN1550
; CURRENT APPLICATION NUMBER: US/09/865,291
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
US-09-865-291-11

Query Match 45.6%; Score 383.6; DB 10; Length 859;
Best Local Similarity 74.2%; Pred. No. 2.2e-92;
Matches 485; Conservative 0; Mismatches 169; Indels 0; Gaps 0;
Qy 4 GTTATCGCTAAACAGATGACCTACAAAGCTTTATATGTCAGGCGCGTCAATGGACACTAC 63
Db 72 GTTATCAAGGAGTTCATGAGGTTTAAGGTCGCTATGGAAGAACGGTCAATGGGACGAG 131
Qy 64 TTTGAGGTCGAAGCGCATGGAAGAAAGGAAAGCTTTACGAGGGGGAGCAGACGGTAAAGGCTG 123
Db 132 TTTGAATAGAAAGCGAAGGAGAGGGAGGCCATACGAAGGCCCAATACCGTAAAGCTT 191
Qy 124 GCTGTACCAAGGGGACCTCTCGCATTTGCTTTGGGATATTTTATCACCACAGTGTGAC 183
Db 192 AAGGTAACCAAGGGGACCTTTGCCATTTGCTTTGGGATATTTTGTCAACCAATTTTCA 251
Qy 184 TACGGAAGCATACATTTCAACCAAGTACCTCGAAGACATCCCTGACTATGTAAGCAGTCA 243
Db 252 TATGGAAGCAAGTATATGTCAAGCACCTCCGACATACCAAGACTATAAAGAGCTGTCA 311
Qy 244 TTCCCGGGGAGATATACATGGGAGGATCATGAACCTTTGAAGATGTTGTCAGTGTGACT 303
Db 312 TTTCTGAAGGATTTAAATGGAAAGGTCATGAACCTTTGAAGACGGTGGCGTCGTTACT 371
Qy 304 GTCAGCAATGATCCAGCATCCAGGCAACTGTTTTCATCTACCATGTCAAGTTCCTCGT 363
Db 372 GTAAACCCAGGATTCAGTTTGCAGATGGCTGTTTTCATCTACAAGGTCAAGTTCATTGCG 431
Qy 364 TTGAACCTTCTCCCAATGGACCTGTTATGCAGAAAGACACAGGGCTGGGAACCCAAAC 423
Db 432 GTGAACCTTCTCCGATGGACCTGTTATGCAGAAAGACAAATGGGCTGGGAAGCCAGC 491
Qy 424 ACTGAGCGTCTCTTTGACAGAGATGGAATGCTGTATAGGAAACAACTTTATGGCTCTGAAG 483
Db 492 ACTGAGCGTTTGTATCTCGTATGGCGGTTTGAAGAGGAGATTCATPAGGCTCTGAAG 551
Qy 484 TTAGAAGAGGTGTGCTACTATTTGTGAATTCAAAATCTACTTACAGGGCAAGAGGCT 543
Db 552 CTGAAGACGGTGTGCTACTATTTGAAATTCAAAAGTATTTTACATGGCAAGAGGCT 611
Qy 544 GTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAAACCAATCAACAAG 603
Db 612 GTGAGCTACAGGGTACTACTATGTTGACTCCAACTGGATATAACAGGCCAACAGAA 671
Qy 604 GATTACACTTCGTTGAGCAGTGTGAATTTCCATTTGCACCGCAACCTGTGGTC 657
Db 672 GACTATACATCGTTGAGCAGTATGAAGAAGACCGGAGGACGCCACCATCTGTTTC 725

RESULT 14

US-10-433-640-12
; Sequence 12, Application US/10433640
; Publication No. US20040115792A1
; GENERAL INFORMATION:
; APPLICANT: Lichtenberg-Frate, Hella
; TITLE OF INVENTION: YEAST STRAIN FOR TESTING THE GENOTOXICITY AND CYTOTOXICITY OF
; FILE REFERENCE: 1487/3
; CURRENT APPLICATION NUMBER: US/10/433,640
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/14610
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: DE 10061872.3
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
US-10-433-640-12

Query Match 45.6%; Score 383.6; DB 17; Length 859;
Best Local Similarity 74.2%; Pred. No. 2.2e-92;
Matches 485; Conservative 0; Mismatches 169; Indels 0; Gaps 0;
Qy 4 GTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGCGCGTCAATGGACACTAC 63
Db 72 GTTATCAAGGAGTTCATGAGGTTTAAGGTCGCTATGGAAGAAAGCGTCAATGGGACGAG 131
Qy 64 TTTGAGGTCGAAGCGCATGGAAGAAAGGAAAGCTTTACGAGGGGGAGCAGACGGTAAAGGCTG 123
Db 132 TTTGAATAGAAAGCGAAGGAGAGGGAGGCCATACGAAGGCCCAATACCGTAAAGCTT 191
Qy 124 GCTGTACCAAGGGGACCTCTCGCATTTGCTTTGGGATATTTTATCACCACAGTGTGAC 183
Db 192 AAGGTAACCAAGGGGACCTTTGCCATTTGCTTTGGGATATTTTGTCAACCAATTTTCA 251
Qy 184 TACGGAAGCATACATTTCAACCAAGTACCTCGAAGACATCCCTGACTATGTAAGCAGTCA 243
Db 252 TATGGAAGCAAGTATATGTCAAGCACCTCCGACATACCAAGACTATAAAGAGCTGTCA 311
Qy 244 TTCCCGGGGAGATATACATGGGAGGATCATGAACCTTTGAAGATGTTGTCAGTGTGACT 303
Db 312 TTTCTGAAGGATTTAAATGGAAAGGTCATGAACCTTTGAAGACGGTGGCGTCGTTACT 371
Qy 304 GTCAGCAATGATCCAGCATCCAGGCAACTGTTTTCATCTACCATGTCAAGTTCCTCGT 363
Db 372 GTAAACCCAGGATTCAGTTTGCAGATGGCTGTTTTCATCTACAAGGTCAAGTTCATTGCG 431
Qy 364 TTGAACCTTCTCCCAATGGACCTGTTATGCAGAAAGACACAGGGCTGGGAACCCAAAC 423
Db 432 GTGAACCTTCTCCGATGGACCTGTTATGCAGAAAGACAAATGGGCTGGGAAGCCAGC 491
Qy 424 ACTGAGCGTCTCTTTGACAGAGATGGAATGCTGTATAGGAAACAACTTTATGGCTCTGAAG 483
Db 492 ACTGAGCGTTTGTATCTCGTATGGCGGTTTGAAGAGGAGATTCATPAGGCTCTGAAG 551
Qy 484 TTAGAAGAGGTGTGCTACTATTTGTGAATTCAAAATCTACTTACAGGGCAAGAGGCT 543
Db 552 CTGAAGACGGTGTGCTACTATTTGAAATTCAAAAGTATTTTACATGGCAAGAGGCT 611
Qy 544 GTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAAACCAATCAACAAG 603
Db 612 GTGAGCTACAGGGTACTACTATGTTGACTCCAACTGGATATAACAGGCCAACAGAA 671
Qy 604 GATTACACTTCGTTGAGCAGTGTGAATTTCCATTTGCACCGCAACCTGTGGTC 657
Db 672 GACTATACATCGTTGAGCAGTATGAAGAAGACCGGAGGACGCCACCATCTGTTTC 725

Search completed: August 14, 2004, 11:00:52
Job time : 466.5 secs

```

RESULT 15
US-09-797-496B-3
; Sequence 3, Application US/09797496B
; Publication NO. US20030049597A1
; GENERAL INFORMATION:
; APPLICANT: Simon, Sanford M.
; APPLICANT: Chen, Yu
; TITLE OF INVENTION: Chimeric Fluorescent Enzymes and Uses Thereof
; FILE REFERENCE: 600-1-267
; CURRENT APPLICATION NUMBER: US/09/797,496B
; CURRENT FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 3311
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Discosoma red fluorescent protein modified as described in specification.
; OTHER INFORMATION: ication.
US-09-797-496B-3

Query Match      45.6%; Score 383.6; DB 10; Length 3311;
Best Local Similarity 74.2%; Pred. No. 4.6e-92;
Matches 485; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

QY      4 GTTATCGCTAAACACATGACCTACAAAGTTTATATGTCAGGCACGGTCAATGGACACTAC 63
Db      307 GTTATCAAGGAGTTTCATGAGGTTTAAGGTTTCGTCATGGAAGGAACGGTCAATGGCACCGAG 366

QY      64 TTTGAGGTCGAAGCGCATGGAAGAAAGCAAGCCCTACGAGGGGGAGCAGACGGTAAAGGCTG 123
Db      367 TTTGAAATAGAGCGGAGGAGGAGGGAGGCCATACGAGGCCCAATACCGTAAAGCTT 426

QY      124 GCTGTCACCAAGGGCGGACCTCTGCCATTGCTGGGATATTTATCACACAGTGTGAC 183
Db      427 AAGGTAACCAAGGGGGACCTTTGCCATTGCTGGGATATTTTGTCAACCAATTTGAC 486

QY      184 TACGGAAGCATACCAATTCACCAAGTACCTGGAAGACATCCCTGACTATGTAAAGCAGTCA 243
Db      487 TATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACCAGACTATAAAAGCTGCA 546

QY      244 TTCCCGGGGAGATATACATGGGAGAGGATCATGAACCTTTGAAGATGGTCAGTGTGACT 303
Db      547 TTTCTGAGAGATTTAATGGGAAGGGTCATGAACCTTTGAAGACGGTGGCGTCGTTACT 606

QY      304 GTCAGCAATGATTCAGCATCCAAAGGCAACGTGTTTCATCTACCATGTCAAGTTCTCTGT 363
Db      607 GTAACCCAGGATTCAGTTTCAGGATGGCTGTTTCATCTACAAGGTCAAGTTCAATTGGC 666

QY      364 TTGAACTTTCTCCCAATGGACCTTTATGCAAGAGAGACACAGGGCTGGGAACCCAAAC 423
Db      667 GTGAACCTTTCTCTCCGATGGACCTTTATGCAAAAGAGACAAATGGGCTGGGAAGCCAGC 726

QY      424 ACTGAGCGCTCTTTGTCAGAGATGGAATGCTGTAGGAAACAACTTTATGCTCTGTAAG 483
Db      727 ACTGAGCGTTTGTATCTCTCGTGTGATGGCGTGTGAAAGGAGAGATTCTAAGGCTCTGAAG 786

QY      484 TTAGAAGAGGTGGTGCATATTTGTGAATTCAAATCTTACTTACAAGGCAAGAAAGCCT 543
Db      787 CTGAAAGACGGTGGTCAATTACCTAGTTGAAATTCAAAGATTTTACATGGCAAGAGCCT 846

QY      544 GTGAAGATGCCAGGATATCATGTGTGACCGCAAACTGGATGTAAACCAATCACAAAG 603
Db      847 GTGCAGCTACCAAGGTACTACTATGTGTGACTCTCAAACTGGATATTAACAAAGCCACACGAA 906

QY      604 GATTACACTTCGTTGAGCAGTGTGAATTTCCATTGTCACGCAACCTGTGTC 657
Db      907 GACTATACAATCGTTGAGCAGTATGAAGAACCCGAGGGACGCCACCATCTGTTC 960
```

This Page Blank (u)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 06:12:47 ; Search time 24.5594 Seconds
(without alignments)
493.990 Million cell updates/sec

Title: US-09-890-463-4
Perfect score: 1287
Sequence: 1 SVIAKQMTYKVMYSGTVNGH.....KPVVACRFRVRKSRHKYAVA 235

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	479.5	37.3	238	US-09-277-716-16	Sequence 16, Appl
2	479.5	37.3	238	US-09-609-161B-16	Sequence 16, Appl
3	479.5	37.3	238	US-09-626-581D-65	Sequence 65, Appl
4	479.5	37.3	238	US-09-415-765B-65	Sequence 65, Appl
5	479.5	37.3	238	US-09-626-580C-65	Sequence 65, Appl
6	470.5	36.6	238	US-09-277-716-32	Sequence 32, Appl
7	470.5	36.6	238	US-09-609-161B-32	Sequence 32, Appl
8	468.5	36.4	238	US-09-839-650-3	Sequence 3, Appl
9	198.5	15.4	238	US-09-472-065A-5	Sequence 5, Appl
10	193.5	15.0	238	US-09-479-645A-12	Sequence 12, Appl
11	192.5	15.0	238	US-09-023-946B-28	Sequence 28, Appl
12	190.5	14.8	238	US-09-023-946B-23	Sequence 23, Appl
13	189.5	14.7	238	US-09-023-946B-36	Sequence 36, Appl
14	189	14.7	238	US-09-023-946B-22	Sequence 22, Appl
15	189	14.7	941	US-09-513-783A-172	Sequence 172, App
16	188.5	14.6	238	US-08-337-915A-2	Sequence 2, Appl
17	188.5	14.6	238	US-08-753-143-2	Sequence 2, Appl
18	188.5	14.6	238	US-08-679-865-2	Sequence 2, Appl
19	188.5	14.6	238	US-08-680-876-2	Sequence 2, Appl
20	188.5	14.6	238	US-08-792-553-2	Sequence 2, Appl
21	188.5	14.6	238	US-08-893-327-16	Sequence 16, Appl
22	188.5	14.6	238	US-08-753-144-2	Sequence 2, Appl
23	188.5	14.6	238	US-09-094-359-2	Sequence 2, Appl
24	188.5	14.6	238	US-09-172-063-1	Sequence 1, Appl
25	188.5	14.6	238	US-09-121-539-1	Sequence 1, Appl
26	188.5	14.6	238	US-09-263-975-2	Sequence 2, Appl
27	188.5	14.6	238	US-08-727-452-2	Sequence 2, Appl

28	188.5	14.6	238	4	US-09-418-785-1	Sequence 1, Appli
29	188.5	14.6	238	4	US-09-214-909-2	Sequence 2, Appli
30	188.5	14.6	238	4	US-09-479-645A-10	Sequence 10, Appl
31	188.5	14.6	238	4	US-09-479-645A-159	Sequence 159, App
32	188.5	14.6	238	4	US-09-129-192C-2	Sequence 2, Appli
33	188.5	14.6	238	4	US-09-129-192C-74	Sequence 74, Appl
34	188.5	14.6	238	4	US-09-602-641-2	Sequence 2, Appli
35	188.5	14.6	238	4	US-09-704-463-2	Sequence 2, Appli
36	188.5	14.6	238	4	US-09-472-085A-4	Sequence 4, Appli
37	188.5	14.6	238	4	US-09-472-085A-6	Sequence 6, Appli
38	188.5	14.6	238	4	US-09-023-946B-29	Sequence 29, Appl
39	188.5	14.6	238	4	US-09-920-922-4	Sequence 4, Appli
40	188.5	14.6	238	5	PCT-US95-14692-2	Sequence 2, Appli
41	188.5	14.6	239	3	US-08-646-538-2	Sequence 2, Appli
42	188.5	14.6	239	3	US-09-094-359-4	Sequence 4, Appli
43	188.5	14.6	239	3	US-09-172-063-3	Sequence 3, Appli
44	188.5	14.6	239	3	US-09-503-222-2	Sequence 2, Appli
45	188.5	14.6	239	4	US-09-513-783A-46	Sequence 46, Appl

ALIGNMENTS

RESULT 1
US-09-277-716-16
; Sequence 16, Application US/092777716A
; Patent No. 6232107
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE
; CURRENT APPLICATION NUMBER: US/09/277,716A
; CURRENT FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: 60/102,939
; EARLIER FILING DATE: 1998-10-01
; EARLIER APPLICATION NUMBER: 60/089,367
; EARLIER FILING DATE: 1998-06-15
; EARLIER APPLICATION NUMBER: 60/079,624
; EARLIER FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla mulleri
; FEATURE:
; OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)
US-09-277-716-16

Query Match	37.3%	Score 479.5;	DB 3;	Length 238;
Best Local Similarity	44.0%	Pred. No. 8.8e-45;		
Matches	96;	Conservative 41;	Mismatches 74;	Indels 7;
Gaps	3;			
QY	7	MTYKMGSTVNGHYFEVGEQKPKYEGEOTVRLAVTKGGPLPFWDLSPQCOYGSIP	66	
Db	15	MSYKNVLEGIYNNHVFTMEGCGKNILFGNQLVQIRVTKGAPLPFAFDIVSPAFYGNRT	74	
QY	67	FTKYPEDIPDYVKQSPGORYTWERIMNFEDGAVCTVSDSSIQGNCFTYHVKFSGLNFP	126	
Db	75	FTKYPNDISDYFIQSFPAGFYERTLRYEDGGLVEIRSDINLIEDKFVRYVEYKGSNFP	134	
QY	127	NGPVMOKTKQGEINTEFLFARDGMILGNFMALKEGGHYLCEFKSYKAKKPVK-MP	185	
Db	135	DGPVMOKTILGIEFSFEAMYNMNGVLGVILVYKLSNGKYYSCHMKTLMSKGVVKEFP	194	
QY	186	GYHYVDKLDVTNNKDYTSVEQCEISAR-----KPV	218	
Db	195	SYHFIQHRLEKT-YVEDGGFVEQHETAQAQWTSIGKPL	231	

RESULT 2
US-09-609-161B-16

```
; Sequence 16, Application US/09609161B
; Patent No. 6436682
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: PROLUME, LTD.
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE LUC
; TITLE OF INVENTION: AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIG
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
; FILE REFERENCE: 24729-121B
; CURRENT APPLICATION NUMBER: US/09/609,161B
; CURRENT FILING DATE: 2000-06-30
; PRIOR FILING DATE: 2000-07-27
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/102,939
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/089,367
; PRIOR FILING DATE: 1998-06-15
; PRIOR APPLICATION NUMBER: 60/079,624
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 16
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla mulleri
; FEATURE:
; OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)
US-09-609-161B-16

Query Match      37.3%; Score 479.5; DB 4; Length 238;
Best Local Similarity 44.0%; Pred. No. 8.8e-45;
Matches 96; Conservative 41; Mismatches 74; Indels 7; Gaps 3;

Qy 7 MTKYKVMGTVNGHYFEVGGKPKYEGEQTIVLAVTKGGPLPFAWDILSPQCOYGSIP 66
Db 15 MSYKVNLEGIYNNHVFVTEGCGKNILFNGQLVQIRVTGAPLPFAFDIVSPAFOYGNRT 74

Qy 67 FTKYPEDIPDYVKOSFPGRYTWERIMNFDGAVCTVSDSIQGNCFIYHVKFSGLNPPP 126
Db 75 FTKYPNDISDYFIQSFPAFGFYERTLYEDGGLVEIRSDINLIEDKFVYRVEYKGSNPPD 134

Qy 127 NGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGHYLCBFKSTYKAKPKVK-MP 185
Db 135 DGPVMQKKTILGIEPSFEAMYNMNGVLGVEVILVYKLSGKYYSCHMKTLMSKGVVKEFP 194

Qy 186 GYHYVDRKLDVTNHNKDYTSVEQCEISTAR-----KPV 218
Db 195 SYHFIQHRLEKT-YVEDGGFVEQHEHETAIAQMTSIGKPL 231

; RESULT 4
US-09-415-765B-65
; Sequence 65, Application US/09415765B
; Patent No. 6548632
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
; TITLE OF INVENTION: Libraries
; FILE REFERENCE: A66900-1/RMS/AMS
; CURRENT APPLICATION NUMBER: US/09/415,765B
; CURRENT FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: 09/169,015
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 65
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla mulleri
US-09-415-765B-65

Query Match      37.3%; Score 479.5; DB 4; Length 238;
Best Local Similarity 44.0%; Pred. No. 8.8e-45;
Matches 96; Conservative 41; Mismatches 74; Indels 7; Gaps 3;

Qy 7 MTKYKVMGTVNGHYFEVGGKPKYEGEQTIVLAVTKGGPLPFAWDILSPQCOYGSIP 66
Db 15 MSYKVNLEGIYNNHVFVTEGCGKNILFNGQLVQIRVTGAPLPFAFDIVSPAFOYGNRT 74

Qy 67 FTKYPEDIPDYVKOSFPGRYTWERIMNFDGAVCTVSDSIQGNCFIYHVKFSGLNPPP 126
Db 75 FTKYPNDISDYFIQSFPAFGFYERTLYEDGGLVEIRSDINLIEDKFVYRVEYKGSNPPD 134

Qy 127 NGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGHYLCBFKSTYKAKPKVK-MP 185
Db 135 DGPVMQKKTILGIEPSFEAMYNMNGVLGVEVILVYKLSGKYYSCHMKTLMSKGVVKEFP 194

Qy 186 GYHYVDRKLDVTNHNKDYTSVEQCEISTAR-----KPV 218
Db 195 SYHFIQHRLEKT-YVEDGGFVEQHEHETAIAQMTSIGKPL 231

; RESULT 3
US-09-626-581D-65
; Sequence 65, Application US/09626581D
; Patent No. 6548249
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
; TITLE OF INVENTION: Libraries
; FILE REFERENCE: A-66900-3/RMS
; CURRENT APPLICATION NUMBER: US/09/626,581D
; CURRENT FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 09/169,015
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 09/415,765
; PRIOR FILING DATE: 1999-10-08
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 65
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla mulleri
US-09-626-581D-65
```

```
Query Match      37.3%; Score 479.5; DB 4; Length 238;
Best Local Similarity 44.0%; Pred. No. 8.8e-45;
Matches 96; Conservative 41; Mismatches 74; Indels 7; Gaps 3;

Qy 7 MTKYKVMGTVNGHYFEVGGKPKYEGEQTIVLAVTKGGPLPFAWDILSPQCOYGSIP 66
Db 15 MSYKVNLEGIYNNHVFVTEGCGKNILFNGQLVQIRVTGAPLPFAFDIVSPAFOYGNRT 74

Qy 67 FTKYPEDIPDYVKOSFPGRYTWERIMNFDGAVCTVSDSIQGNCFIYHVKFSGLNPPP 126
Db 75 FTKYPNDISDYFIQSFPAFGFYERTLYEDGGLVEIRSDINLIEDKFVYRVEYKGSNPPD 134

Qy 127 NGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGHYLCBFKSTYKAKPKVK-MP 185
Db 135 DGPVMQKKTILGIEPSFEAMYNMNGVLGVEVILVYKLSGKYYSCHMKTLMSKGVVKEFP 194

Qy 186 GYHYVDRKLDVTNHNKDYTSVEQCEISTAR-----KPV 218
Db 195 SYHFIQHRLEKT-YVEDGGFVEQHEHETAIAQMTSIGKPL 231

; RESULT 4
US-09-415-765B-65
; Sequence 65, Application US/09415765B
; Patent No. 6548632
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
; TITLE OF INVENTION: Libraries
; FILE REFERENCE: A66900-1/RMS/AMS
; CURRENT APPLICATION NUMBER: US/09/415,765B
; CURRENT FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: 09/169,015
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 65
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla mulleri
US-09-415-765B-65

Query Match      37.3%; Score 479.5; DB 4; Length 238;
Best Local Similarity 44.0%; Pred. No. 8.8e-45;
Matches 96; Conservative 41; Mismatches 74; Indels 7; Gaps 3;

Qy 7 MTKYKVMGTVNGHYFEVGGKPKYEGEQTIVLAVTKGGPLPFAWDILSPQCOYGSIP 66
Db 15 MSYKVNLEGIYNNHVFVTEGCGKNILFNGQLVQIRVTGAPLPFAFDIVSPAFOYGNRT 74

Qy 67 FTKYPEDIPDYVKOSFPGRYTWERIMNFDGAVCTVSDSIQGNCFIYHVKFSGLNPPP 126
Db 75 FTKYPNDISDYFIQSFPAFGFYERTLYEDGGLVEIRSDINLIEDKFVYRVEYKGSNPPD 134

Qy 127 NGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGHYLCBFKSTYKAKPKVK-MP 185
Db 135 DGPVMQKKTILGIEPSFEAMYNMNGVLGVEVILVYKLSGKYYSCHMKTLMSKGVVKEFP 194

Qy 186 GYHYVDRKLDVTNHNKDYTSVEQCEISTAR-----KPV 218
Db 195 SYHFIQHRLEKT-YVEDGGFVEQHEHETAIAQMTSIGKPL 231

; RESULT 5
US-09-626-580C-65
; Sequence 65, Application US/09626580C
; Patent No. 6562617
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Peele, Beau R.
; TITLE OF INVENTION: FUSIONS OF SCAFFOLD PROTEINS WITH RANDOM PEPTIDE LIBRARIES
```



```
; FILE REFERENCE: 25436/1755
; CURRENT APPLICATION NUMBER: US/09/839,650
; CURRENT FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 3
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla muelleri
US-09-839-650-3

Query Match      36.4%; Score 468.5; DB 4; Length 238;
Best Local Similarity 43.6%; Pred. No. 1.4e-43;
Matches 95; Conservative 40; Mismatches 76; Indels 7; Gaps 3;

Qy 7 MTKYVMSGTNGHYVEEGDGKPYEGEQTIVRLAVTKGGPLPAWDILSPQCOYGSIP 66
Db 15 MSYKVNLEGIYNNHVTMEGCKRNILFGNQLVHIVTKGGPLPAFADIVSPAFQYGNRT 74

Qy 67 FTKYPEDIPYVKQSFGRYTWERIMNPDGAVCTVSNDSSIQNCFIYHVKFSGLNPP 126
Db 75 FTKYPNDISDYFIQSFPAGFYERTLRYEDGGLVEIRSDINLIEDKFVYRVEYKGSNPD 134

Qy 127 NGPVMQKTOGWENTERLRFARDGMLIGNNFMALKLEGGHYLCEFFKTYKAKXPVK-MP 185
Db 135 DGPVMQKTIIGIEPSFEAMVWNGVLGVEVILVYKLSGKYYSCHMTLMKSGVVKRFP 194

Qy 186 GYHVVDKLDVTNHNKDYTSVEQCEISTAR-----KPV 218
Db 195 SYFHQHRLEKT-YVEDGGFVEQHETATAQWTSIGKPL 231

RESULT 9
US-09-472-065A-5
; Sequence 5, Application US/09472065A
; Patent No. 6638732
; GENERAL INFORMATION:
; APPLICANT: Evans, Krista
; TITLE OF INVENTION: Mutants of Green Fluorescent Protein
; CURRENT APPLICATION NUMBER: US/09/472,065A
; CURRENT FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: US 09/970,762
; PRIOR FILING DATE: 1997-11-14
; PRIOR APPLICATION NUMBER: US 60/030,935
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 5
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Aequorea victoria, A1 mutant
US-09-472-065A-5

Query Match      15.4%; Score 198.5; DB 4; Length 238;
Best Local Similarity 25.7%; Pred. No. 9.2e-14;
Matches 53; Conservative 46; Mismatches 90; Indels 17; Gaps 8;

Qy 11 VYMSGTVNGHYVEEGDGKPYEGEQTIVRLAVTKGGPLPAWDILSPQCOYGSIPFTKY 70
Db 16 VELDGDVNGHKFSVSGEGGATYKGLTKFICTT-GKLPVWPVTLVTTCAYGVCFSRY 74

Qy 71 PEDIP--DYVKQSFGRYTWERIMNPDGAVCTVSNDSSIQNCFIYHVKFSGLNPPNG 128
Db 75 PDHMKQHDFFKSAMPEGYVQERTIFFKDDGNKYKTRAEVKFEGDITLVNRIELKIDFKEDG 134

Qy 129 PVMQKKTQGWEPNTERLF-----ARDGMLIGNNFMALK-LEGGHYLCEFF--KSTYKAKK 180
Db 135 NILGHKLE-YNNSHNVYIMADKQNGIKV--NFKIRNIHEDGSQLADHYQONTPIGDG 191

181 PVKMPGYHYVDKLDVT---NHNKDY 203
192 PVLIPDNHYLTSQALSCKDPNEKRHD 217

US-09-472-065A-5

Query Match      15.4%; Score 198.5; DB 4; Length 238;
Best Local Similarity 25.7%; Pred. No. 9.2e-14;
Matches 53; Conservative 46; Mismatches 90; Indels 17; Gaps 8;

Qy 11 VYMSGTVNGHYVEEGDGKPYEGEQTIVRLAVTKGGPLPAWDILSPQCOYGSIPFTKY 70
Db 16 VELDGDVNGHKFSVSGEGGATYKGLTKFICTT-GKLPVWPVTLVTTCAYGVCFSRY 74

Qy 71 PEDIP--DYVKQSFGRYTWERIMNPDGAVCTVSNDSSIQNCFIYHVKFSGLNPPNG 128
Db 75 PDHMKQHDFFKSAMPEGYVQERTIFFKDDGNKYKTRAEVKFEGDITLVNRIELKIDFKEDG 134

Qy 129 PVMQKKTQGWEPNTERLF-----ARDGMLIGNNFMALK-LEGGHYLCEFF--KSTYKAKK 180
Db 135 NILGHKLE-YNNSHNVYIMADKQNGIKV--NFKIRNIHEDGSQLADHYQONTPIGDG 191

181 PVKMPGYHYVDKLDVT---NHNKDY 203
192 PVLIPDNHYLTSQALSCKDPNEKRHD 217

US-09-472-065A-5

Query Match      15.4%; Score 198.5; DB 4; Length 238;
Best Local Similarity 25.7%; Pred. No. 9.2e-14;
Matches 53; Conservative 46; Mismatches 90; Indels 17; Gaps 8;

Qy 11 VYMSGTVNGHYVEEGDGKPYEGEQTIVRLAVTKGGPLPAWDILSPQCOYGSIPFTKY 70
Db 16 VELDGDVNGHKFSVSGEGGATYKGLTKFICTT-GKLPVWPVTLVTTCAYGVCFSRY 74

Qy 71 PEDIP--DYVKQSFGRYTWERIMNPDGAVCTVSNDSSIQNCFIYHVKFSGLNPPNG 128
Db 75 PDHMKQHDFFKSAMPEGYVQERTIFFKDDGNKYKTRAEVKFEGDITLVNRIELKIDFKEDG 134

Qy 129 PVMQKKTQGWEPNTERLF-----ARDGMLIGNNFMALK-LEGGHYLCEFF--KSTYKAKK 180
Db 135 NILGHKLE-YNNSHNVYIMADKQNGIKV--NFKIRNIHEDGSQLADHYQONTPIGDG 191

181 PVKMPGYHYVDKLDVT---NHNKDY 203
192 PVLIPDNHYLTSQALSCKDPNEKRHD 217

US-09-472-065A-5

Query Match      15.0%; Score 193.5; DB 4; Length 238;
Best Local Similarity 25.2%; Pred. No. 3.3e-13;
Matches 52; Conservative 47; Mismatches 90; Indels 17; Gaps 8;

Qy 11 VYMSGTVNGHYVEEGDGKPYEGEQTIVRLAVTKGGPLPAWDILSPQCOYGSIPFTKY 70
Db 16 VELDGDVNGHKFSVSGEGGATYKGLTKFICTT-GKLPVWPVTLVTTCAYGVCFSRY 74

Qy 71 PEDIP--DYVKQSFGRYTWERIMNPDGAVCTVSNDSSIQNCFIYHVKFSGLNPPNG 128
Db 75 PDHMKQHDFFKSAMPEGYVQERTIFFKDDGNKYKTRAEVKFEGDITLVNRIELKIDFKEDG 134

Qy 129 PVMQKKTQGWEPNTERLF-----ARDGMLIGNNFMALK-LEGGHYLCEFF--KSTYKAKK 180
Db 135 NILGHKLE-YNNSHNVYIMADKQNGIKV--NFKIRNIHEDGSQLADHYQONTPIGDG 191

181 PVKMPGYHYVDKLDVT---NHNKDY 203
192 PVLIPDNHYLTSQALSCKDPNEKRHD 217

US-09-479-645A-12
; Sequence 12, Application US/09479645A
; Patent No. 6489141
; GENERAL INFORMATION:
; APPLICANT: FRAZER, Ian Hector
; APPLICANT: ZHOU, Jian
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND METHOD FOR SELECTIVELY
; FILE REFERENCE: 210338.0001/IUS
; CURRENT APPLICATION NUMBER: US/09/479,645A
; CURRENT FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: PCT/AU98/00530
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: AU P07765
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: AU P09467
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 12
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic gfp
; OTHER INFORMATION: gene (Papillomavirus)
; FEATURE:
; OTHER INFORMATION: Codons of humanized gfp gene replaced with
; OTHER INFORMATION: synonymous codons used at relatively high
; OTHER INFORMATION: frequency by papillomavirus genes
US-09-479-645A-12

Query Match      15.0%; Score 193.5; DB 4; Length 238;
Best Local Similarity 25.2%; Pred. No. 3.3e-13;
Matches 52; Conservative 47; Mismatches 90; Indels 17; Gaps 8;

Qy 11 VYMSGTVNGHYVEEGDGKPYEGEQTIVRLAVTKGGPLPAWDILSPQCOYGSIPFTKY 70
Db 16 VELDGDVNGHKFSVSGEGGATYKGLTKFICTT-GKLPVWPVTLVTTCAYGVCFSRY 74

Qy 71 PEDIP--DYVKQSFGRYTWERIMNPDGAVCTVSNDSSIQNCFIYHVKFSGLNPPNG 128
Db 75 PDHMKQHDFFKSAMPEGYVQERTIFFKDDGNKYKTRAEVKFEGDITLVNRIELKIDFKEDG 134

Qy 129 PVMQKKTQGWEPNTERLF-----ARDGMLIGNNFMALK-LEGGHYLCEFF--KSTYKAKK 180
Db 135 NILGHKLE-YNNSHNVYIMADKQNGIKV--NFKIRNIHEDGSQLADHYQONTPIGDG 191

181 PVKMPGYHYVDKLDVT---NHNKDY 203
192 PVLIPDNHYLTSQALSCKDPNEKRHD 217

US-09-023-946B-28
; Sequence 28, Application US/09023946B
; Patent No. 6670449
; GENERAL INFORMATION:
; APPLICANT: GERO MIESENBOCK, ET AL.
; TITLE OF INVENTION: HYBRID MOLECULES AND
; CELLULAR MICROENVIRONMENTS
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
```

```

CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/023,946B
  FILING DATE: 13-Feb-1998
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 60/036,805
  FILING DATE: 14 FEBRUARY 1997
  APPLICATION NUMBER: 60/038,179
  FILING DATE: 13 FEBRUARY 1997
ATTORNEY/AGENT INFORMATION:
  NAME: KENNETH H. SONNENSELD
  REGISTRATION NUMBER: 33,285
  REFERENCE/DOCKET NUMBER: 2955-4004US2
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (212) 758-4800
  TELEFAX: (212) 751-6849
  TELEX: 421792
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
  LENGTH: 238
  TYPE: AMINO ACID
  STRANDEDNESS: UNKNOWN
  TOPOLOGY: UNKNOWN
SEQUENCE DESCRIPTION: SEQ ID NO: 23:
US-09-023-946B-23

Query Match      14.8%; Score 190.5; DB 4; Length 238;
Best Local Similarity 26.2%; Pred. No. 7e-13;
Matches 54; Conservative 44; Mismatches 91; Indels 17; Gaps 8;

Qy 11 VYMSGTVNGHYFEVGGDKGKPYEGEQTVLAVTKGGPLPFAMDLSPQCOYGSIPFTKY 70
Db 16 VELDGDVNGHKFSVSGEGEDATYGLTKLFICTT-GKLPVPMPTLVTTTSYGVCFSRY 74
Qy 71 PEDIP--DYUKQFPFGRYTWERIMNPDGAVCTVSNDSISQGNCFIYHVKESGLNFPNG 128
Db 75 PDHWKRHDFFSKAMPEGYVQERTIFFKDDGNKYTRAEVKPEGDTLVNRIELKGIKDFKDG 134
Qy 129 PVWOKTKQGWEPNTERLF-----ARDQMLGNPF-MALKEGGGHYLCEF--KSTYKAKK 180
Db 135 NILGHKLE-YVNDHQVYIMADQKNG--IKANFKIRHNIEDGGVQLADHYQQTPTGDG 191
Qy 181 PVKMPGYHYVDRKLDVT---NHNKDY 203
Db 192 PVLLEDNHYLHTQSALSKDPNEKRDH 217

RESULT 13
US-09-023-946B-36
; Sequence 36, Application US/09023946B
; Patent No. 6670449
; GENERAL INFORMATION:
; APPLICANT: GERO MIESENBOCK, ET AL.
; TITLE OF INVENTION: HYBRID MOLECULES AND
; THEIR USE FOR OPTICALLY DETECTING CHANGES IN
; CELLULAR MICROENVIRONMENTS
;
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,946B
; FILING DATE: 13-Feb-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/036,805

```


QY 181 PVKMEGYHYVDRKLDVT---NHNKDY-----TSVE 207
| | | | | : : : : :
Db 193 FVLLPDNHYLSTQSALS KDPNEKR DDMVLLFEFVTAAGITLGNDELYKSGMSVVGIDLGFQ 252
QY 208 QCEISIR-----KPVVACRPFRRVKSRHKYAVA 235
| | | | | : : : : :
Db 253 SCYVAVARAGGIETIANEYSDRCTPACISFGPKNRSIGAAA 293

Search completed: August 12, 2004, 06:21:07
Job time : 25.5594 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2004, 11:19:50 ; Search time 2783.5 Seconds
(without alignments)
9022.487 Million cell updates/sec

Title: US-09-890-463-5
Perfect score: 841
Sequence: 1 tccgtatcgttaaacagat.....aaaagcgcgcgtcgaatta 841

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	58.6	7.0	866	14	CK159216 FGAS04062
C 2	55.6	6.6	667	29	CNS02Y2V
C 3	54.6	6.5	1065	29	AL220324 Tetraodon
C 4	54.4	6.5	297	14	AL106553 Drosophil
					CF805258 lad65f03.

5	54.4	6.5	820	13	BU563278
6	54.4	6.5	1201	9	AL514477
C 7	53.8	6.4	1201	13	EX398622
8	53.4	6.3	359	29	CNS007CK
9	53.4	6.3	362	14	CB958074
10	53.4	6.3	752	13	BU567098
C 11	53.4	6.3	861	10	BE895801
12	53.2	6.3	504	14	CF513671
13	53.2	6.3	675	14	CD641892
14	53.2	6.3	1201	13	EX395452
C 15	53	6.3	539	13	EX403488
16	53	6.3	1201	13	EX446698
17	53	6.3	1237	14	CF661002
18	52.8	6.3	278	13	BO566832
19	52.8	6.3	422	9	AU262401
C 20	52.8	6.3	867	29	CNS0054A
C 21	52.8	6.3	998	13	EX453291
22	52.6	6.3	225	14	CF550581
23	52.6	6.3	371	29	CNS0460V
C 24	52.6	6.3	417	13	EX465145
25	52.6	6.3	918	9	AL558886
26	52.4	6.2	812	12	BM985610
C 27	52.4	6.2	855	14	CK197555
28	52.2	6.2	300	12	BI397139
29	52.2	6.2	314	10	BE047863
30	52.2	6.2	417	10	BF294370
C 31	52.2	6.2	553	13	EX446453
32	52.2	6.2	926	29	CNS028L5
C 33	52.2	6.2	1184	13	EX446507
C 34	52	6.2	239	14	CB981957
C 35	52	6.2	784	9	AL048397
C 36	52	6.2	855	14	CK196558
C 37	52	6.2	884	14	CA946208
38	52	6.2	1201	9	AL559982
39	51.8	6.2	308	12	BI537549
40	51.8	6.2	318	13	BU716755
41	51.8	6.2	528	14	CA324119
42	51.8	6.2	579	13	EX415932
43	51.8	6.2	683	12	BI527017
44	51.8	6.2	779	28	AZ175769
45	51.8	6.2	1095	14	CF661586

ALIGNMENTS

RESULT 1
CK159216/c
LOCUS CK159216 866 bp mRNA linear EST 05-DEC-2003
DEFINITION FGAS040623 Triticum aestivum FGAS: TaLT5 Triticum aestivum cDNA,
mRNA sequence.
ACCESSION CK159216
VERSION CK159216.1 GI:38985155
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
1 (bases 1 to 866)
Allard, F., Crosby, M.L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D.,
Genswein, B., Graf, R., Gulick, P., Hrycan, L.D., Laroche, A.,
Links, M.G., McCarthy, E.L., Monroy, A., Muzak, I., Nilsson, D.,
Penniket, C., Roach, J.L. and Sarhan, F.
Functional Genomics of Abiotic Stress in Wheat and Canola Crops
Unpublished (2003)
Contact: Wm L Crosby
Bioinformatics
University of Saskatchewan, Department of Computer Science
1C101 Engineering Building, 57 Campus Drive, Saskatoon,
Saskatchewan, S7N 5A9, Canada
Tel: 306 966 1769
Fax: 306 966 2033

Email: fgas_ests@cs.usask.ca
 This sequence is the direct result of the Base calling software
 Phred (default parameters). It is the raw base calls. To aid in the
 identification of the high quality insert the software Lucy
 (default parameters) has been run on this sequence. Lucy identified
 the region [124,600].
 Plate: Ta1t538 row: B column: 11.

FEATURES

Location/Qualifiers
 1..866
 /organism="Triticum aestivum"
 /mol_type="mRNA"
 /cultivar="Wheat line PI 178383"
 /db_xref="taxon:4565"
 /lab_host="DH5 alpha"
 /clone_lib="Triticum aestivum FGAS: Ta1t5"
 /note="Organ: Crown; Vector: pGEM-T; SSH (suppression
 subtractive hybridization) cDNA library from genotype
 PI178383 cold hardened at 2 C for 21 days and 49 days
 (equal amount of cDNA pooled together before subtraction,
 tester) and subtracted against genotype Norstar cold
 hardened at 2 C for 1 day (24 H)(driver). Modified Smart
 cDNA (Clontech) priming and non-directional cloning"

ORIGIN

Query Match 7.0%; Score 58.6; DB 14; Length 866;
 Best Local Similarity 61.4%; Pred. No. 3.4;
 Matches 94; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
 Qy 673 AGAGTCAATCAAGGCACAAATAGCAGTCGCGTAAACAACTGATTCGATTTAGCT 732
 Db 384 AAAATAAAACACAGAAAAATTTTGGTTTAAAGCAAAAAAGGGGTTTTTAATTTTGT 325
 Qy 733 TAGAGAGTAGGACGACGAAGTGTAGACACCTTCATGATTAAACCTTTTGAACCAAC 792
 Db 324 AAAAAAATAAACTGGGAAAAATATAAAAAATTAACACGGAATTTTTTAAAAAA 265
 Qy 793 SCCAAAAAATAAAAAAATAAAAAAATAAAAAA 825
 Db 264 AAAAAAATAAAAAAATAAAAAAATAAAAAA 232

RESULT 2

CNS02YZV/c
 LOCUS 667 bp DNA linear GSS 01-SEP-2000
 DEFINITION Tetraodon nigroviridis genome survey sequence PUC-ORI end of clone
 18N15 of library G from Tetraodon nigroviridis, genomic survey
 sequence.

ACCESSION

AL220324
 VERSION AL220324.1 GI:7879143
 KEYWORDS GSS; genome survey sequence.
 SOURCE Tetraodon nigroviridis
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodontidae; Tetraodon.

REFERENCE

1
 Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 Estimate of human gene number provided by genome-wide analysis
 using Tetraodon nigroviridis DNA sequence
 Nat. Genet. 25 (2), 235-238 (2000)

JOURNAL

MEDLINE 20296633
 PUBMED 10835645
 2
 Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
 Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
 Saurin,W., Bernot,A. and Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis

TITLE

JOURNAL Genome Res. 10 (7), 939-949 (2000)
 MEDLINE 20359837

PUBMED

10899143

REFERENCE

3 (bases 1 to 667)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
 - Web : www.genoscope.cns.fr)
 This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.

FEATURES

source

Location/Qualifiers
 1..667
 /organism="Tetraodon nigroviridis"
 /mol_type="genomic DNA"
 /db_xref="taxon:99883"
 /clone="181N15"
 /clone_lib="G"
 /note="Genoscope sequence ID : CONG181CG08SP1-end :
 PUC-ORI"

ORIGIN

Query Match 6.6%; Score 55.6; DB 29; Length 667;
 Best Local Similarity 48.1%; Pred. No. 12;
 Matches 76; Conservative 24; Mismatches 58; Indels 0; Gaps 0;
 Qy 678 CAATCAAGGCACAAATAAGCAGTCGCGTAAACAACTGATTCGATTTAGCTTAGAG 737
 Db 201 CWAACACAAARAAAAAANWRTTAAAAAATAAAAAAATAAAAAAATAAAAAA 142
 Qy 738 AAGTAGAACGACGAAGTGTAGACACCTTCAATGATTAACTTTGAAACACACCCAA 797
 Db 141 AAAAAAARMACMAMMAAFAACACAAATAATCAACAAACAAAAAATAAAAAA 82
 Qy 798 AAAAAAATAAAAAAATAAAAAAATAAAAAAAGCGCGCTC 835
 Db 81 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAACAGCTCTC 44

RESULT 3

CNS016BZ/c

LOCUS

DEFINITION

CNS016BZ 1065 bp DNA linear GSS 26-JUL-1999
 Drosophila melanogaster genome survey sequence T7 end of BAC
 BACN15B05 of DrosBAC library from Drosophila melanogaster (fruit
 fly), genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT


```
Db 99 TTGTAATGAGATCTGATGCGCTCTTCTTGGAGTGTCTGAAGACAGCTACAGTGTATCTCTA 158
Qy 730 GCTTAGAGAGTACGACGAGAGTGTAGACACCTTCATGATTAACTTTTGAAAC 789
Db 159 TAATAAATAATTAATCTTTAAATAAATAAATAAATAAATAAATAAATAAATAA 218
Qy 790 AACSCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 825
Db 219 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 254

RESULT 6
AL514477/c 1201 bp mRNA linear EST 08-MAY-2003
LOCUS
DEFINITION AL514477 Homo sapiens NEUROBLASTOMA Homo sapiens cDNA clone
CLOBB003ZA09 3-PRIME, mRNA sequence.
ACCESSION AL514477
VERSION AL514477.2 GI:30464362
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:12777971.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4172.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CLOBB003ZA09FPI&cluster=4172.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CLOBB003ZA09FPI.

FEATURES
Location/Qualifiers
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CLOBB003ZA09"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 6.5%; Score 54.4; DB 9; Length 1201;
Best Local Similarity 46.7%; Pred. No. 14;
Matches 98; Conservative 20; Mismatches 92; Indels 0; Gaps 0;

Qy 628 GAATATTCATTCAGCGAACCCTTTGGTCGCCCTGCTTTTTCAGAGTCAATCAGG 687
Db 281 GGACTTACCTTTTCATTTGAAMADGCGCATACCTAGATTATTTTGGKAAAAA 222
Qy 688 CACAAATAGCAGTGGCGTAAACAGCTAGATTCTGATTTCAGCTAGAGAGTAGGAAC 747
Db 221 AGTAAATTTTAAATCTTAARGGKRXKADATATATTAATAGTAAATGTATG 162
Qy 748 GAAGAAGTGTAGACACCTTCATGATTAAACTTTTGAAACACSCCAAAAAAAAA 807
Db 161 TAATMAAATAACAMTMTTATGTTTATTAATAATGAAAAAATAAATAAATAA 102
Qy 808 AAAAAAAAAAAAAAAAAAAGCGCCGCTCGA 837
```

```
Db 101 AAAAAAAAAAAAAAAAAAGGCGCGCTCTA 72

RESULT 7
BX398622/c 1201 bp mRNA linear EST 13-MAY-2003
LOCUS
DEFINITION BX398622 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI061YJ01 5-PRIME, mRNA sequence.
ACCESSION BX398622
VERSION BX398622.1 GI:30617636
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CSODI061CE01QPI.

FEATURES
Location/Qualifiers
source
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI061YJ01"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 6.4%; Score 53.8; DB 13; Length 1201;
Best Local Similarity 46.7%; Pred. No. 17;
Matches 64; Conservative 28; Mismatches 45; Indels 0; Gaps 0;

Qy 689 ACRAATAGCAGTGGCGTAAACAGCTAGATTCTGATTTCAGTATAGAGAGTAGGACG 748
Db 437 AAAAAAGAGAAAAAARAAAAATAATATTTTWTWWTTTWWATWWAAAAAAW 378
Qy 749 AAGAAGTGTAGACACCTTCATGATTAACTTTGAAACACSCCAAAAAAAAA 808
Db 377 TWWAWWAAWAAAAATTTWWWWWWWWWWAAAAAATAAATAAATAAATAA 318
Qy 809 AAAAAAAAAAAAAAAAAAAAA 825
Db 317 AAAAAAAAAAAAAAAAAAAAA 301

RESULT 8
CNS007CK 359 bp DNA linear GSS 03-JUN-1999
LOCUS
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR15J06 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL066995
VERSION AL066995.1 GI:4945559
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30351140"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_184"
/notes="Organ: Pooled-Glandular; Vector: pDNR-LIB; Site_1: SfiI (ggcattatgcc); Site 2: SfiI (ggcgctgcgc); Library is oligo-dT primed and directionally cloned. cDNA was prepared from a glandular pool of tissues from thyroid, parathyroid, adrenal, cortex and pineal gland. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGCGGGCAGCATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.38 kb (range 0.60-3.5 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match          6.3%; Score 53.4; DB 14; Length 362;
Best Local Similarity 56.8%; Pred.No. 35;
Matches 99; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY      651   TTTCGTGCGCTGCTGTTTTTTTCAGAGTCAAAATCAAGSCACAAAATAAGCAGTGCGGTAAAA 710
        |||    |||    |||    |||    |||    |||    |||    |||    |||    |||
Db       137   TTTATTACTTTGTCITTAATAATCAATGCAGAGAAGTTGTTGACTGTAGGGGAAT 196

QY      711   AACGTAGATCTGATTTTAGCTTTAGAGAAGTAGGAACGAGAGTGTAGACAACCTTCAA 770
        |||    |||    |||    |||    |||    |||    |||    |||    |||    |||
Db       197   AAAGTTAATTCAAATTTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 256

QY      771   TGATTAACTTTGAAACACSCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 825
        |||    |||    |||    |||    |||    |||    |||    |||    |||    |||
Db       257   AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 311


RESULT 10
BU567098
LOCUS           BU567098              752 bp     mRNA         linear     EST 16-SEP-2002
DEFINITION     AGENCOURT 10393772 NIH_MGC 141 Homo sapiens cDNA clone
IMAGE:6606668 5', mRNA_sequence.
ACCESSION     BU567098
VERSION       BU567098.1 GI:22917398
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1 (bases 1 to 752)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs@mail.nih.gov
             Tissue Procurement: NCBI
             cDNA Library Preparation: Michael Brownstein Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCW2852 row: k column: 20
             High quality sequence stop: 400.
             Location/Qualifiers
               1..752
                 /organism="Homo sapiens"
                 /mol_type="mRNA"
                 /db_xref="taxon:9606"
                 /clone="IMAGE:660668"
                 /isolate_type="mixed (pool of 40 RNAs)"
                 /lab_host="NIH_MGC_184"

FEATURES
source

```

/clone_lib="NIH_MGC_141"
/notes="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);
Site 2: SfiI (ggccctctggcc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were
used in cloning as follows:
5'-AAGCTTAGCGTATCAACGAGATGGCTACGGCGGG-3' and
5'-ATTCTAGACGCGCGGCCGACATG-TT(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 0.2-0.5
kb size fraction (other fractions present in NIH MGC 142).
Library created in the laboratory of M. Brownstein (NIH,
NIH). Note: this is a NIH MGC library."

```

ORIGIN
Query Match      6.3%; Score 53.4; DB 13; Length 752;
Best Local Similarity 56.6%; Pred. No. 25;
Matches          99; Conservative    0; Mismatches 76; Indels   0; Gaps   0;

QY      651 TTGTGGTCGCTGCTGTTTTTTCACAGTCACAATCAAGGCACACAATAAGCAGCTGGCGGTAAAA 710
        ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db       203 TTTATTTTACTTGTTGTCCTTTAAATTC AATTCAGAGAA GTTGT TGACTGTAGCGGGAAT 262

QY      711 AACGTAGATCTCATTTTAGCTTAGAAGAAGT AGCAAGCAAGTAAGTGTAGACAACTTCAA 770
        ||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db       263 AAAGTTAATTC AATTTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 322

QY      771 TGATTAACCTTTTGAAAACACSCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 825
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db       323 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 377


RESULT 11
BB895801/c
LOCUS
DEFINITION
BB895801                861 bp     mRNA           linear      EST 20-OCT-2000
601432727F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3917935 5',
mRNA sequence.

```

Site_2: Salt; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life Technologies."

ORIGIN

Query Match 6.3%; Score 53.4; DB 10; Length 861;
Best Local Similarity 56.6%; Pred. No. 23;
Matches 99; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 651 TTGTGTCGCGTGTGTTTTTTCAGATCAAAATCAAGGCAAAATAGCAGTGCCTAAAA 710
Db 224 TTTTACTCACCCCTGTTTGTGGAACAAAAACAATGAATAGATAAAATAAAAA 165
QY 711 AACGTAGATCTGATTTTAGCTTAGAGAAGTGAAGACGAAGTGTAGACAACTTCAA 770
Db 164 AAAGAAAAACGCTTTTGTGTGAGAAAGACGACCCCAAAAAA 105
QY 771 TGATTAACCTTTTGAAACCAACSCCAAAAAA 825
Db 104 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 50

RESULT 12
CF513671
LOCUS
DEFINITION
CF513671 504 bp mRNA linear EST 09-SEP-2003
Cabud0007 IIF_D11 Vitis vinifera cv. cabernet sauvignon (Clone 8)
Bud - CABUD Vitis vinifera cDNA clone Cabud0007_IIF_D11_5', mRNA
sequence.
CF513671
CF513671.1 GI:34545439
EST.
Vitis vinifera
Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; Vitaceae; Vitis.
1 (bases 1 to 504)
Goes da Silva,F., Iandolino,A., Lim,H., Baek,J., Jones,K. and
Cook,P.
Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
berries at various developmental stages
Unpublished (2003)
Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcook@ucdavis.edu
Seq primer: ACGGTACCGGACATATGCC.
Location/Qualifiers
1..504
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon (Clone 8)"
/db_xref="taxon:29760"
/clone="Cabud0007_IIF_D11"
/sex="Hermaphrodite"
/dev_stage="Pre-bloom (10-11 days before bloom)"
/lab_host="DH5alpha"
/clone_lib="Vitis vinifera cv. cabernet sauvignon (Clone
8) Bud - CABUD"
/notes="Organ: Bud; Vector: pDNR; Site_1: Sfil; Site_2:
Sfil; CABUD is a cDNA library of Vitis vinifera cv.
'Cabernet Sauvignon' Clone 8 dissected buds. Samples were
collected May 13, 2002 from pre-bloom plants (10-11 days
before bloom), pre-veraison. Sampled vines were located at
the University of California, Davis. Experimental
vineyard. cDNAs were made by oligo-dT priming and
directionally cloned. 5' and 3' adaptors were used in
cloning as follows:
5'-AAGCATGGTATCAACGACAGTGCCTATCGCGGG-3' and
5'-ATTCTAGGCGGCGGCGGCGGACATG-DT (30)NN-3' Library was

FEATURES
source

constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

Query Match 6.3%; Score 53.2; DB 14; Length 504;
Best Local Similarity 51.2%; Pred. No. 32;
Matches 124; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 585 TGTACCAATCACAAAGATTACATTCCTGAGCAGCGTGAATTTCCATTGCGAG 644
|||||
Db 226 TGTACATAAATATAGGATACATTCCTTTCTCCATCAATCTGTTTATCC 285
|||||

QY 645 CAAACCTTGGCGCTGCTGTTTTTTCAGAGTCAAAATCAAGGCACAAATAGCAGTGGC 704
|||||
Db 286 TAAACATTTGGTGAACATACTCTTTTCGCGAGTCGAATACAACTCAAAAAA 345
|||||

QY 705 GTAAAAACGTTAGATCTGATTTTAGCTTTAGAGTAGTAGGACGAGAGTGTCAGACAC 764
|||||
Db 346 AAAAAA 405
|||||

QY 765 CTTCAATGATTAACTTTTGAACACACSCCAAAAAA 824
|||||

Db 406 AAAAAA 465

QY 825 AG 826
||

Db 466 AG 467

RESULT 13

CD641892

LOCUS

DEFINITION CD641892 675 bp mRNA linear EST 17-JUN-2003

AGENCOURT 14535240 NIH MGC 191 Homo sapiens cDNA clone

IMAGE:30415374 5', mRNA sequence.

ACCESSION CD641892

VERSION CD641892

KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE

AUTHORS

TITLE NIH-MGC <http://mgc.nci.nih.gov/>.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgapbs-remail.nih.gov

Tissue Procurement: Narayan Bhat

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>

Plate: NDCM213 row: h column: 07

High quality sequence stop: 312.

FEATURES

source

1. .675

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30415374"

/tissue_type="Pooled"

/lab_host="DH10B (T1 phage-resistant)"

/clone_lib="NIH MGC 191"

/note="Vector: pONR-LIB; Site 1: SfiI (ggccattatggcc); Site 2: SfiI (ggcgctcgcc); Library is oligo-dT primed and directionally cloned. FBMC - Peripheral Blood Mononuclear Cells. RNA was pooled from 3/6hour stimulation with PMA adn Ionomycin. 5' and 3' adaptors were used in

cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGAGCGCCGACATG-dr(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.69 kb (range 0.70-5.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 6.3%; Score 53.2; DB 14; Length 675;
Best Local Similarity 57.1%; Pred. No. 28;
Matches 97; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 656 TCGCTGCTGTTTTTTTTCAGAGTCAAAATCAAGGCACAAATAGCAGTGGGTAAAAACGT 715
|||||

Db 111 TCACCTGCTGTTTTTTTATATAAAATGTGTACAAAGTTAATTTATTCATTAAATAAGC 170
|||||

QY 716 AGATTCTGATTTTGTAGAGTAGTGAAGAGTGTAGACAACTTCAATGATT 775
|||||

Db 171 TCTTTAACTATAAAAAA 230
|||||

QY 776 AAACCTTTGAAAAACACSCCAAAAAA 825
|||||

Db 231 AAAAAA 280
|||||

RESULT 14

BX395452

LOCUS

DEFINITION BX395452 1201 bp mRNA linear EST 13-MAY-2003

CDNA clone CS0DD008VH18 5-PRIME, mRNA sequence.

ACCESSION BX395452

VERSION BX395452.1 GI:30612818

KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens (human)

REFERENCE

AUTHORS

TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

JOURNAL Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

COMMENT Full-length cDNA libraries and normalization Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: secref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6286.f For more information about this cluster, see <http://www.genoscope.cns.fr/>

<http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DD008DD09QPl&cluster=6286.f>. Contact : Feng Liang Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DD008DD09QPl.

FEATURES

source

1. 1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DD008VH18"

/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"

/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 6.3%; Score 53.2; DB 13; Length 1201;
Best Local Similarity 52.1%; Pred. No. 22;
Matches 74; Conservative 16; Mismatches 52; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2004, 11:19:50 ; Search time 409.5 seconds
(without alignments)
8724.627 Million cell updates/sec

Title: US-09-890-463-5

Perfect score: 841

Sequence: 1 tccgtatcgtctaaacagat.....aaaagcgccgcgtcgaaatta 841

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Jan04.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002s.*
- 7: geneseqn2003as.*
- 8: geneseqn2003bs.*
- 9: geneseqn2003cs.*
- 10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	840.6	100.0	841	3	AAA52082 Pigment p
2	824.6	98.0	841	3	Aaa52083 Pigment p
3	651.4	77.5	693	6	Abz12089 Colour Fa
4	650.4	77.3	660	6	Abz12084 Colour Fa
5	648.8	77.1	660	6	Abz12085 Colour Fa
6	648.8	77.1	660	6	Abz12074 Colour Fa
7	648.8	77.1	660	6	Abz12069 Colour Fa
8	648.8	77.1	669	6	Abz12156 Colour Fa
9	647.6	77.0	663	6	Abz12070 Colour Fa
10	647.2	77.0	660	6	Abz12072 Colour Fa
11	640.8	76.2	660	6	Abz12083 Colour Fa
12	637.6	75.8	660	6	Abz12098 Colour Fa
13	637.6	75.8	669	6	Abz12155 Colour Fa
14	636	75.6	669	6	Abz12161 Colour Fa
15	636	75.6	669	6	Abz12158 Colour Fa
16	636	75.6	669	6	Abz12157 Colour Fa
17	634.4	75.4	669	6	Abz12159 Colour Fa
18	633	75.3	663	6	Abz12129 Colour Fa
19	633	75.3	663	6	Abz12128 Colour Fa
20	631.4	75.1	663	6	Abz12107 Colour Fa
21	631.4	75.1	663	6	Abz12106 Colour Fa
22	631.4	75.1	765	6	Abz12103 Colour Fa
23	631.4	75.1	765	6	Abz12102 Colour Fa

24	628.4	74.7	660	6	Abz12104 Colour Fa
25	626.6	74.5	660	6	Abz12105 Colour Fa
26	626.4	74.5	660	6	Abz12100 Colour Fa
27	626.4	74.5	660	6	Abz12079 Colour Fa
28	626.4	74.5	660	6	Abz12078 Colour Fa
29	626.4	74.5	660	6	Abz12094 Colour Fa
30	626.4	74.5	660	6	Abz12081 Colour Fa
31	626.4	74.5	660	6	Abz12101 Colour Fa
32	625	74.3	663	6	Abz12110 Colour Fa
33	625	74.3	663	6	Abz12113 Colour Fa
34	625	74.3	663	6	Abz12116 Colour Fa
35	624.8	74.3	660	6	Abz12082 Colour Fa
36	623.4	74.1	660	6	Abz12142 Colour Fa
37	623.4	74.1	663	6	Abz12132 Colour Fa
38	623.4	74.1	663	6	Abz12108 Colour Fa
39	623.4	74.1	663	6	Abz12109 Colour Fa
40	623.2	74.1	660	6	Abz12080 Colour Fa
41	623.2	74.1	660	6	Abz12097 Colour Fa
42	623.2	74.1	660	6	Abz12095 Colour Fa
43	623	74.1	660	6	Abz12077 Colour Fa
44	622.6	74.0	660	6	Abz12075 Colour Fa
45	621.8	73.9	660	6	Abz12135 Colour Fa

ALIGNMENTS

RESULT 1

AAA52082
ID AAA52082 standard; cdna; 841 BP.
XX
AC AAA52082;
XX
DT 04-DEC-2000 (first entry)
XX
DE Pigment protein from coral tissue POC3 cdna.
XX
KW N-terminal; pigment protein from coral tissue; POC3; fluorescence;
KW tissue marker; fluorescent marker; dyestuff; sunscreen; ultra violet;
KW UV filter; POC3; ss.
XX
OS Acropora aspera.
XX
FH Key Location/Qualifiers
FT CDS 1..696
FT /*tag= a
FT /label= POC3
FT /product= "Pigment_protein_from_coral_tissue"
FT /partial
XX
WO2000046233-A1.
XX
PD 10-AUG-2000.
XX
PF 02-FEB-2000; 2000WO-AU0000056.
XX
PR 02-FEB-1999; 99AU-00008463.
XX
(UNSY) UNIV SYDNEY.
XX
PI Hoegh-Guldberg O, Dove S;
XX
DR WPI: 2000-532892/48.
DR P-PSDB; AAY97149.
XX
PT Novel pigment protein derived from corals capable of emitting
PT fluorescence upon irradiation by incident light useful as tissue marker,
XX fluorescent marker or general dyestuff.
XX
PS Claim 10; Page 44; 49pp; English.
XX
CC cdna libraries were constructed from a blue pigmented coral, Acropora
CC aspera to isolate sequences encoding polypeptides with N-terminal

CC sequences as shown in AAY97147-48. Pigment protein from coral tissue
CC (PPCT) is capable of emitting fluorescence upon irradiation by incident
CC light whose maximal absorbance is in the range of 320-600 nm and a
CC maximal fluorescence emission is in the range of 300-700 nm. PPCT may be
CC used as a tissue marker, fluorescent marker (e.g. to follow gene
CC expression in transformed tissues) or general dyestuff (all claimed).
CC PPCT may also be used in sunscreen formulations or UV filters (both
CC claimed)
XX
SQ Sequence 841 BP; 274 A; 171 C; 196 G; 199 T; 0 U; 1 Other;
Query Match 100.0%; Score 840.6; DB 3; Length 841;
Best Local Similarity 100.0%; Pred. No. 1.2e-196;
Matches 841; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGCGGTCAATGGACAC 60
DB 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGCGGTCAATGGACAC 60
QY 61 TACTTTGAGTTCGAAGGCGATGGAAGAAAGAAAGCTTACGAGGGGAGCAGCGTTAAGG 120
DB 61 TACTTTGAGTTCGAAGGCGATGGAAGAAAGAAAGCTTACGAGGGGAGCAGCGTTAAGG 120
QY 121 CTGGCTGTCAACAAGGGCGACCTCTGCCATTGCTTGGGATATTTATACCAAGTGT 180
DB 121 CTGGCTGTCAACAAGGGCGACCTCTGCCATTGCTTGGGATATTTATACCAAGTGT 180
QY 181 CAGTACGGAAGCATACCAATTCACCAAGTACCTGAGAGATCCCTGACTATGTAAGCAG 240
DB 181 CAGTACGGAAGCATACCAATTCACCAAGTACCTGAGAGATCCCTGACTATGTAAGCAG 240
QY 241 TCATTCCCGGGAGATATACATGAGGAGGATCATGAATCTTGAAGATGTCAGTGTGT 300
DB 241 TCATTCCCGGGAGATATACATGAGGAGGATCATGAATCTTGAAGATGTCAGTGTGT 300
QY 301 ACTGTCAGCAATGATTCAGCATCCAGGCAACTGTTTTCATCTACCATGTCAGTGTCT 360
DB 301 ACTGTCAGCAATGATTCAGCATCCAGGCAACTGTTTTCATCTACCATGTCAGTGTCT 360
QY 361 GGTGTTGAATCTTCTCCCAATGGACCTGTATGCAAGAGAGACACAGGCTGGGAACC 420
DB 361 GGTGTTGAATCTTCTCCCAATGGACCTGTATGCAAGAGAGACACAGGCTGGGAACC 420
QY 421 AACACTGAGCGTCTCTTTGCAAGATGGAATGCTGATAGGAAACAACTTTATGCTCTG 480
DB 421 AACACTGAGCGTCTCTTTGCAAGATGGAATGCTGATAGGAAACAACTTTATGCTCTG 480
QY 481 AAGTTAGAGAGGTGTCATATTTGTGTAATCAATCTTACATCAAGGCAAGGAG 540
DB 481 AAGTTAGAGAGGTGTCATATTTGTGTAATCAATCTTACATCAAGGCAAGGAG 540
QY 541 CTTGTGAAGATGCCAGGCTATCCTATGTTGACCGCAAACTGGATGTAACCAATCACAAC 600
DB 541 CTTGTGAAGATGCCAGGCTATCCTATGTTGACCGCAAACTGGATGTAACCAATCACAAC 600
QY 601 AAGGATTACACTTCGGTTGAGCAGCGTGAAATTTCCATTGCAAGCAAACTTTGTCGCC 660
DB 601 AAGGATTACACTTCGGTTGAGCAGCGTGAAATTTCCATTGCAAGCAAACTTTGTCGCC 660
QY 661 TGCTGTTTTTTCAGAGTCAATCAAGGCAAAATTAAGCAGTGGCGTAAACCGTAGATT 720
DB 661 TGCTGTTTTTTCAGAGTCAATCAAGGCAAAATTAAGCAGTGGCGTAAACCGTAGATT 720
QY 721 CTGATTTTAGCTTAGAGAGTAGGAAAGAGAGTGTAGAACCTTCAATGATTAAGT 780
DB 721 CTGATTTTAGCTTAGAGAGTAGGAAAGAGAGTGTAGAACCTTCAATGATTAAGT 780
QY 781 TTTGAAAAACAACSCCAAAAAAATAAAAAAATAAAAAAATAAAAAAAGCGCGCTCGAATT 840
DB 781 TTTGAAAAACAACSCCAAAAAAATAAAAAAATAAAAAAATAAAAAAAGCGCGCTCGAATT 840
QY 841 A 841

DB 841 A 841
RESULT 2
ID AAA52083 standard; cDNA; 841 BP.
XX
AC AAA52083;
XX
DT 04-DEC-2000 (first entry)
XX
DE Pigment protein from coral tissue POC4 cDNA.
XX
KW N-terminal; pigment protein from coral tissue; PPCT; fluorescence;
KW tissue marker; fluorescent marker; dyestuff; sunscreen; ultra violet;
XX UV filter; POC3; ss.
XX
OS Acropora aspera.
XX
FH Key Location/Qualifiers
CDS 1..708
FT /*tag= a
FT /label= POC4
FT /product= "Pigment protein from coral tissue"
FT /partial
PN WO200046233-A1.
XX
PD 10-AUG-2000.
XX
PF 02-FEB-2000; 2000WO-AU0000056.
XX
PR 02-FEB-1999; 99AU-00008463.
XX (UNSY) UNIV SYDNEY.
XX
PI Hoegh-Guldberg O, Dove S;
XX
XX WPI; 2000-532892/48.
DR P-PSDB; AAY97150.
XX
XX Novel pigment protein derived from corals capable of emitting
XX fluorescence upon irradiation by incident light useful as tissue marker,
XX fluorescent marker or general dyestuff.
XX
PS Claim 10; Page 44; 49pp; English.
XX
CC cDNA libraries were constructed from a blue pigmented coral, Acropora
CC aspera to isolate sequences encoding polypeptides with N-terminal
CC sequences as shown in AAY97147-48. Pigment protein from coral tissue
CC (PPCT) is capable of emitting fluorescence upon irradiation by incident
CC light whose maximal absorbance is in the range of 320-600 nm and a
CC maximal fluorescence emission is in the range of 300-700 nm. PPCT may be
CC used as a tissue marker, fluorescent marker (e.g. to follow gene
CC expression in transformed tissues) or general dyestuff (all claimed).
CC PPCT may also be used in sunscreen formulations or UV filters (both
CC claimed)
SQ Sequence 841 BP; 275 A; 171 C; 195 G; 200 T; 0 U; 0 Other;
Query Match 98.0%; Score 824.6; DB 3; Length 841;
Best Local Similarity 98.7%; Pred. No. 9.9e-193;
Matches 830; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
QY 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGCGGTCAATGGACAC 60
DB 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGCGGTCAATGGACAC 60
QY 61 TACTTTGAGTTCGAAGGCGATGGAAGAAAGAAAGCTTACGAGGGGAGCAGCGTTAAGG 120
DB 61 TACTTTGAGTTCGAAGGCGATGGAAGAAAGAAAGCTTACGAGGGGAGCAGCGTTAAGG 120
QY 121 CTGGCTGTCTCACCAGGCGGACCTCTGCCATTGCTTGGGATATTTATACCAAGTGT 180

Db 121 CTGGCTGTCAACAGGGCGACCTCTGCCATTTGCTTGGATATTTATCACCAAGTGT 180
Qy 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAGCAG 240
Db 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAGCAG 240
Qy 241 TCATTCCCGGGAGATATACATGCGAGAGATCATGAACCTTTGAAGATGGTCAAGTGTGT 300
Db 241 TCATTCCCGGGAGATATACATGCGAGAGATCATGAACCTTTGAAGATGGTCAAGTGTGT 300
Qy 301 ACTGTACGAATGATTCAGCATCCAAAGCAACTGTTTCATCTACCATGTCAAGTTCTCT 360
Db 301 ACTGTACGAATGATTCAGCATCCAAAGCAACTGTTTCATCTACCATGTCAAGTTCTCT 360
Qy 361 GATTGGAACCTTCTCCCAATGAGTATGATGAGAGATGATGATGAGAGAGAGAGAGAG 420
Db 361 GATTGGAACCTTCTCCCAATGAGTATGATGAGAGATGATGATGAGAGAGAGAGAGAG 420
Qy 421 AACACTGAGCGTCTCTTTCACGAGATGAATGCTGATAGGAAACAACTTTATGGCTGTG 480
Db 421 AACACTGAGCGTCTCTTTCACGAGATGAATGCTGATAGGAAACAACTTTATGGCTGTG 480
Qy 481 AAGTTAGAGAGAGTGGTCACTATTGTTGTAATTCATCTTACCAAGGCAAGGAG 540
Db 481 AAGTTAGAGAGAGTGGTCACTATTGTTGTAATTCATCTTACCAAGGCAAGGAG 540
Qy 541 CCGTGCAGATCCCGGATCACTATGTTGACCGCAACTGGATGATACCAATCAAC 600
Db 541 CCGTGCAGATCCCGGATCACTATGTTGACCGCAACTGGATGATACCAATCAAC 600
Qy 601 AAGGATTACACTTCCGTTGAGCAGCGTGAATTTCCATTGCACGCAACCTTTTGGTCGC 660
Db 601 AAGGATTACACTTCCGTTGAGCAGCGTGAATTTCCATTGCACGCAACCTTTTGGTCGC 660
Qy 661 TGCTGTTTTTTCAGAGTCAAAATCAAGGCAAAATAGCAGTGGCGGTAAACACGTAGATT 720
Db 661 TGCCCGTTTTTTCAGAGTCAAAATCAAGGCAAAATAGCAGTGGCGGTAAACACGTAGATT 720
Qy 721 CTGATTTAGCTTAGAAGTAGGACGAGAGAGTGTAGACAACTTCAATGATTAAC 780
Db 721 CTGATTTAGCTTTAGAGTAGGACGAGAGAGTGTAAACCAACCATTAATGATTAAC 780
Qy 781 TTTGAAAACACACCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGAATT 840
Db 781 TTTGAAAACACGCGCATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGCGCGCTCGAATT 840
Qy 841 A 841
Db 841 A 841

RESULT 3

ABZ12089

ID ABZ12089 standard; DNA; 693 BP.

XX AC

XX ABZ12089;

XX

DT 22-JAN-2003 (first entry)

XX DE

XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 61.

XX

XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;

XX KW

XX chromophore; biomatrix; transgenic animal; colouring agent;

XX KW

XX flower industry; expression marker; reporter molecule; photon trap;

XX UV sink; sunscreen; ds.

XX OS

XX Porites murrayensis.

XX FN

XX WO200270703-A2.

XX

XX 12-SEP-2002.

XX

PF 01-MAR-2002; 2002WO-CB0000928.
XX
PR 02-MAR-2001; 2001US-0273227P.
PR 21-MAR-2001; 2001AU-0000387A.
PR 15-OCT-2001; 2001US-0329816P.
XX
PA (NUFA-) NUFARM LTD.
PA (YQYU) UNIV QUEENSLAND.
PA (JONE/) JONES E L.
XX
PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG,
PI Hoegh-Guldberg IO, Prescott M;
XX
XX WPI; 2002-740765/80.
XX
PT Novel color-facilitating molecule for producing a biomatrix, has a
PT polypeptide which alone/along with molecules imparts altered visual
PT characteristics to cells in the absence of excitation by extraneous non-
PT white light.

Claim 6; Page 338-339; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM) comprising a polypeptide which, in a cell, alone or together with one or more other molecules imparts an altered visual characteristic to the cell when visualised by a human eye in the absence of excitation by extraneous non-white light or particle emission. CFMs are useful for producing a transgenic animal which exhibits a novel colour e.g. sheep with blue or red coloured fleece. They are useful for producing coloured plant extracts, e.g. flavouring, beverage or juice or colouring agent. Other uses include transducing or intensifying an image, providing additional light for growing phototropic organisms e.g. algae and/or corals, for coating materials that experience UV damage e.g. plastics and car upholstery. CFMs are useful in the flower industry, in the development of new varieties of flowering plants. Other contemplated uses include, expression markers, general reporter molecules, photon traps, UV sinks or in sunscreens. CFMs modify visible colour in edible and/or ornamental fungal species, and in fruits and vegetables to enhance their marketability. CFMs embedded in a gel matrix improve image quality in situations of distorted light spectra (biomatrix). The first all-protein chromophore to be isolated was Green Fluorescent protein (GFP). The sequences given in records ABZ12068-ABZ12195 represent CFM related DNA sequences

XX SQ Sequence 693 BP; 204 A; 152 C; 167 G; 170 T; 0 U; 0 Other;

Query Match 77.5%; Score 651.4; DB 6; Length 693;

Best Local Similarity 96.2%; Pred. No. 3.4e-150;

Matches 667; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATGTCAGGACGGTCAATGGACAC 60
Db 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATGTCAGGACGGTCAATGGACAC 60
Qy 61 TACTTTGAGTTCGAAGGCGATGGAAGAGAGAGCTTACGAGGGGAGCAGCGTAAAG 120
Db 61 TACTTTGAGTTCGAAGGCGATGGAAGAGAGAGCTTACGAGGGGAGCAGCGTAAAG 120
Qy 121 CTGGCTGTCAACAGGGCGGACCTCTGCCATTTGCTTGGGATATTTATCACCAAGTGT 180
Db 121 CTCACTGTCAACAGGGCGGACCTCTGCCATTTGCTTGGGATATTTATCACCAAGTGT 180
Qy 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAGCAG 240
Db 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCGAAGACATCCCTGACTATGTAAGCAG 240
Qy 241 TCATTCCCGGGAGATATACATGCGAGAGATCATGAACCTTTGAAGATGGTCAAGTGTGT 300
Db 241 TCATTCCCGGGAGATATACATGCGAGAGATCATGAACCTTTGAAGATGGTCAAGTGTGT 300
Qy 301 ACTGTACGAATGATTCAGCATCCAAAGCAACTGTTTCATCTACCATGTCAAGTTCTCT 360
Db 301 ACTGTACGAATGATTCAGCATCCAAAGCAACTGTTTCATCTACCATGTCAAGTTCTCT 360

361 GGTGTTGAACCTTTCTCCCAATGGACCTTTATGAGAGAGAGACACACAGGGCTGGGAACCC 420
|||||
361 GGTGTTGAACCTTTCTCCCAATGGACCTTTATGAGAGAGAGACACACAGGGCTGGGAACCC 420
|||||
421 AACACTGAGCGTCTCTTTGACGAGAGATGCTGATAGGAAACAACTTTATGGCTCTG 480
|||||
421 AACACTGAGCGTCTCTTTGACGAGAGATGCTGATAGGAAACAACTTTATGGCTCTG 480
|||||
481 AAGTTAGAGAGAGGTGCTACATTTTGTGTAATCAATCTACTTACAGGCAAGGAAG 540
|||||
481 AAGTTAGAGAGAGGTGCTACATTTTGTGTAATCAATCTACTTACAGGCAAGGAAG 540
|||||
541 CCTGTGAAGATGCCAGGCTATCCTATCTGTGACCGCAAACTGGATTAACCAATCACAAAC 600
|||||
541 CCTGTGATGATGCCAGGCTATCCTATCTGTGACCGCAAACTGGATTAACCAATCACAAAC 600
|||||
601 AAGGATTACACTTCCGTTGAGCAGCGTGAATTTCCATTTGCACTGACGCAAACTTTGCTGCC 660
|||||
661 TGCTGTTTTCAGAGTCAATCAAGGCACAAA 693
|||||
661 TGACGTTTTTTCAGAGTCAATCAAGGCACAAA 693
|||||

RESULT 4
ABZ12084
ID ABZ12084 standard; DNA; 660 BP.
XX
AC ABZ12084;
XX
DT 22-JAN-2003 (first entry)
XX
DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 51.
XX
KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
KW UV sink; sunscreen; ds.
XX
OS Millepora sp.
XX
PN WO200270703-A2.
XX
PD 12-SEP-2002.
XX
PF 01-MAR-2002; 2002WO-GB000928.
XX
PR 02-MAR-2001; 2001US-0273227P.
PR 21-MAR-2001; 2001AU-00003874.
PR 15-OCT-2001; 2001US-0329816P.
XX
XX (NUFA-) NUFARM LTD.
PA (UYOU) UNIV QUEENSLAND.
PA (JONE/) JONES E L.
XX
PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
PI Hoegh-Guldberg IO, Prescott M;
XX
DR WPI; 2002-740765/80.
XX
PT Novel color-facilitating molecule for producing a biomatrix, has a
PT polypeptide which alone/along with molecules imparts altered visual
PT characteristics to cells in the absence of excitation by extraneous non-
PT white light.
XX
PS Claim 6; Page 326-327; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)
CC comprising a polypeptide which, in a cell, alone or together with one or
CC more other molecules imparts an altered visual characteristic to the cell
CC when visualised by a human eye in the absence of excitation by extraneous

CC non-white light or particle emission. CFMs are useful for producing a
CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
CC red coloured fleece. They are useful for producing coloured plant
CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
CC uses include transducing or intensifying an image, providing additional
CC light for growing phototropic organisms e.g. algae and/or corals, for
CC coating materials that experience UV damage e.g. plastics and car
CC upholstery. CFMs are useful in the flower industry, in the development of
CC new varieties of flowering plants. Other contemplated uses include,
CC expression markers, general reporter molecules, photon traps, UV sinks or
CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
CC fungal species, and in fruits and vegetables to enhance their
CC marketability. CFMs embedded in a gel matrix improve image quality in
CC situations of distorted light spectra (biomatrix). The first all-protein
CC chromophore to be isolated was Green Fluorescent protein (GFP). The
CC sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
CC sequences

XX
SQ Sequence 660 BP; 192 A; 146 C; 163 G; 159 T; 0 U; 0 Other;
Query Match 77.3%; Score 650.4; DB 6; Length 660;
Best Local Similarity 99.1%; Pred. No. 6e-150;
Matches 554; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGAGGACCGTCAATGGACAC 60
DB 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGAGGACCGTCAATGGACAC 60
QY 61 TACTTTGAGTTCGAGGCGGATGGAAGAAAGCCCTTACGAGGGGAGCAGCGTAAAGG 120
DB 61 TACTTTGAGTTCGAGGCGGATGGAAGAAAGCCCTTACGAGGGGAGCAGCGTAAAGG 120
QY 121 CTGGCTGTCAACCAAGGCGGACCTCTGCCATTTTGTCTGGGATATTTTATCACCAAGTGT 180
DB 121 CTGACTGTCAACCAAGGCGGACCTCTGCCATTTTGTCTGGGATATTTTATCACCAAGTGT 180
QY 181 CAGTACGGAAGCATACCATTTACCAAGTACCTTGAAGCATCCCTGACTATGTAAGCAG 240
DB 181 CAGTACGGAAGCATACCATTTACCAAGTACCTTGAAGCATCCCTGACTATGTAAGCAG 240
QY 241 TCATTTCCCGGAGGATATACATGGGAGGATCATGAATTTGAAGATGGTGCAGTGTGT 300
DB 241 TCATTTCCCGGAGGATATACATGGGAGGATCATGAATTTGAAGATGGTGCAGTGTGT 300
QY 301 ACTGTGAGCAATGATTCCAGCATCCCAAGGCAACTGTTTCACTACCATGTCAAGTTCTCT 360
DB 301 ACTGTGAGCAATGATTCCAGCATCCCAAGGCAACTGTTTCACTACCATGTCAAGTTCTCT 360
QY 361 GGTGTTGAACCTTTCTCCCAATGGACCTGTTATGAGAGAGACACACAGGGCTGGGAACCC 420
DB 361 GGTGTTGAACCTTTCTCCCAATGGACCTGTTATGAGAGAGACACACAGGGCTGGGAACCC 420
QY 421 AACACTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG 480
DB 421 AACACTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG 480
QY 481 AAGTTAGAGAGAGGTGCTACATTTTGTGTAATCAATCTACTTACAGGCAAGGAAG 540
DB 481 AAGTTAGAGAGAGGTGCTACATTTTGTGTAATCAATCTACTTACAGGCAAGGAAG 540
QY 541 CCTGTGAGATGCCAGGCTATCCTATCTGTGACCGCAAACTGGATTAACCAATCACAAAC 600
DB 541 CCTGTGAGATGCCAGGCTATCCTATCTGTGACCGCAAACTGGATTAACCAATCACAAAC 600
QY 601 AAGGATTACACTTCCGTTGAGCAGCGTGAATTTCCATTTGCACTGACGCAAACTTTGCTGCC 660
DB 601 AAGGATTACACTTCCGTTGAGCAGCGTGAATTTCCATTTGCACTGACGCAAACTTTGCTGCC 660

RESULT 5
ABZ12085
ID ABZ12085 standard; DNA; 660 BP.
XX

AC ABZ12085;
 DT 22-JAN-2003 (first entry)
 XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 53.
 DE
 XX Colour facilitating molecule; CPM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunscreen; ds.
 XX
 OS Millepora sp.
 XX
 XX WO200270703-A2.
 FN
 PD 12-SEP-2002.
 XX
 XX 01-MAR-2002; 2002WO-GB000928.
 PF
 XX 02-MAR-2001; 2001US-0273227P.
 PR
 PR 21-MAR-2001; 2001AU-00003874.
 PR
 PR 15-OCT-2001; 2001US-0329816P.
 XX
 XX (NUFA-) NUFARM LTD.
 FA (UYQU) UNIV QUEENSLAND.
 FA (JONE/) JONES E L.
 XX
 XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
 PI Hoesgh-Guldberg IO, Prescott M;
 XX
 DR WPI; 2002-740765/80.
 XX
 XX Novel color-facilitating molecule for producing a biomatrix, has a
 PT polypeptide which alone/along with molecules imparts altered visual
 PT characteristics to cells in the absence of excitation by extraneous non-
 PT white light.
 XX
 PS Claim 6; Page 329; 510pp; English.
 XX
 XX The invention relates to an isolated colour-facilitating molecule (CFM)
 CC comprising a polypeptide which, in a cell, alone or together with one or
 CC more other molecules imparts an altered visual characteristic to the cell
 CC more other molecules imparts an altered visual characteristic to the cell
 CC non-whitelight by a human eye in the absence of excitation by extraneous
 CC non-whitelight or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant
 CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development of
 CC new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
 CC sequences
 XX
 SQ Sequence 660 BP; 192 A; 147 C; 163 G; 158 T; 0 U; 0 Other;
 Query Match 77.1%; Score 648.8; DB 6; Length 660;
 Best Local Similarity 98.9%; Pred. No. 1.5e-14;
 Matches 653; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGTCAGGACGGTCAATGGACAC 60
 DB 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGTCAGGACGGTCAATGGACAC 60
 QY 61 TACITTTAGGTCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGGAGCAGCGTAAGG 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 61 TACITTTAGGTCGAAGGCGATGGAAGAAAGCCCTTACGAGGGGGAGCAGCGTAAGG 120
 QY 121 CTGGCTGTCAACAAGGGCGGACCTCTGCCATTGTGGATATTTTATCACACAGTGT 180
 Db 121 CTGACTGTCAACAAGGGCGGACCTCTGCCATTGTGGATATTTTATCACACAGTCA 180
 QY 181 CAGTACGGAAGCATACCAATTCCCAAGTACCTCGAAGCATCCCTGACTATGTAAGCAG 240
 Db 181 CAGTACGGAAGCATACCAATTCCCAAGTACCTCGAAGCATCCCTGACTATGTAAGCAG 240
 QY 241 TCATTTCCCGGGGAGATATACATGGGAGAGGATCATGAACCTTTGAAGATGGTGCAGTGT 300
 Db 241 TCATTTCCCGGGGAGATATACATGGGAGAGGATCATGAACCTTTGAAGATGGTGCAGTGT 300
 QY 301 ACTGTCAGCAATGANTCCAGCATCCCAAGCAACTGTCTTCACTACCATGTCAAGTTCCT 360
 Db 301 ACTGTCAGCAATGANTCCAGCATCCCAAGCAACTGTCTTCACTACCATGTCAAGTTCCT 360
 QY 361 GGTTTGAACCTTTCTCCCAATGGACCTGTATGCAGAGAGAGACACAGGGCTGGGAACC 420
 Db 361 GGTTTGAACCTTTCTCCCAATGGACCTGTATGCAGAGAGAGACACAGGGCTGGGAACC 420
 QY 421 AACCTGAGCGTCTCTTTGACGAGATGGAATGCTGTATAGGAAACAACTTTATGGCTCTG 480
 Db 421 AACCTGAGCGTCTCTTTGACGAGATGGAATGCTGTATAGGAAACAACTTTATGGCTCTG 480
 QY 481 AAGTTAGAAGGAGGTGGTCACTATTGTTGTGAATTCAAATCTACTTACAGGCAAGGAAG 540
 Db 481 AAGTTAGAAGGAGGTGGTCACTATTGTTGTGAATTCAAATCTACTTACAGGCAAGGAAG 540
 QY 541 CCTGTGAAGATGCCAGGGTATCATCTATGTCACCGCAACTGGATGTAACCAATCACAAAC 600
 Db 541 CCTGTGAAGATGCCAGGGTATCATCTATGTCACCGCAACTGGATGTAACCAATCACAAAC 600
 QY 601 AAGGATTACACTTCGGTTGAGCAGCGTCAAAATTTCCATTGCACGCAAACTTTGGTGGCC 660
 Db 601 AAGGATTACACTTCGGTTGAGCAGCGTCAAAATTTCCATTGCACGCAAACTTTGGTGGCC 660
 RESULT 6
 ID ABZ12074 standard; DNA; 660 BP.
 XX
 AC ABZ12074;
 XX
 DT 07-AUG-2003 (revised)
 DT 22-JAN-2003 (first entry)
 XX
 DT Colour Facilitating molecule (CFM) related sequence #SEQ ID 31.
 DE
 DE Colour facilitating molecule; CPM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunscreen; ds.
 XX
 OS Caulastrea sp.
 XX
 XX WO200270703-A2.
 FN
 PD 12-SEP-2002.
 XX
 XX 01-MAR-2002; 2002WO-GB000928.
 XX
 PR 02-MAR-2001; 2001US-0273227P.
 PR 21-MAR-2001; 2001AU-00003874.
 PR 15-OCT-2001; 2001US-0329816P.
 XX
 XX (NUFA-) NUFARM LTD.
 FA (UYQU) UNIV QUEENSLAND.
 FA (JONE/) JONES E L.
 XX
 XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
 PI Hoesgh-Guldberg IO, Prescott M;

XX	WPI; 2002-740765/80.	
DR	Novel color-facilitating molecule for producing a biomatrix, has a	
XX	polypeptide which alone/along with molecules imparts altered visual	
PT	characteristics to cells in the absence of excitation by extraneous non-	
PT	white light.	
PT		
XX	Claim 6; Page 297-298; 510pp; English.	
XX	The invention relates to an isolated colour-facilitating molecule (CFM)	
CC	comprising a polypeptide which, in a cell, alone or together with one or	
CC	more other molecules imparts an altered visual characteristic to the cell	
CC	when visualised by a human eye in the absence of excitation by extraneous	
CC	non-white light or particle emission. CFMs are useful for producing a	
CC	transgenic animal which exhibits a novel colour e.g. sheep with blue or	
CC	red coloured fleece. They are useful for producing coloured plant	
CC	extracts, e.g. flavouring, beverage or juice or colouring agent. Other	
CC	uses include transducing or intensifying an image, providing additional	
CC	light for growing phototropic organisms e.g. algae and/or corals, for	
CC	coating materials that experience UV damage e.g. plastics and car	
CC	upholstery. CFMs are useful in the flower industry, in the development of	
CC	new varieties of flowering plants. Other contemplated uses include,	
CC	expression markers, general reporter molecules, photon traps, UV sinks or	
CC	in sunscreens. CFMs modify visible colour in edible and/or ornamental	
CC	in sunscreens. CFMs modify visible colour in edible and/or ornamental	
CC	funeral species, and in fruits and vegetables to enhance their	
CC	marketability. CFMs embedded in a gel matrix improve image quality in	
CC	situations of distorted light spectra (biomatrix). The first all-protein	
CC	chromophore to be isolated was Green Fluorescent protein (GFP). The	
CC	sequences given in records ABZ12068-ABZ12195 represent CFM related DNA	
CC	sequences. (Updated on 07-AUG-2003 to correct OS field.)	
XX	Sequence 660 BP; 192 A; 147 C; 163 G; 158 T; 0 U; 0 Other;	
SQ	Query Match 77.1%; Score 648.8; DB 6; Length 660;	
	Best Local Similarity 98.9%; Pred. No. 1.5e-149;	
	Matches 653; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
QY	1 TCCGTTATCGCTAACAGATGACCTACAAAGTATATATGTCAGGACGCTCAATGGACAC 60	
DB	1 TCCGTTATCGCTAACAGATGACCTACAAAGTATATATGTCAGGACGCTCAATGGACAC 60	
QY	61 TACTTTGAGGTCGAGGCGATGGAAGGAAAGCCCTTACGAGGGGGAGCAGCGTAAAG 120	
DB	61 TACTTTGAGGTCGAGGCGATGGAAGGAAAGCCCTTACGAGGGGGAGCAGCGTAAAG 120	
QY	121 CTGGCTGTACCAAGGGGCGACCTCTGCCATTTGCTTGGGATATTTATACCAAGTGT 180	
DB	121 CTGACTGTACCAAGGGGCGACCTCTGCCATTTGCTTGGGATATTTATACCAAGTGT 180	
QY	181 CAGTACGGAAGCATACCATTCACCAAGTACCCCTGAAGACATCCCTGACTATGTAAGCAG 240	
DB	181 CAGTACGGAAGCATACCATTCACCAAGTACCCCTGAAGACATCCCTGACTATGTAAGCAG 240	
QY	241 TCATTCCCGGGAGATATACATGGAGAGGATCATGAACCTTTGAAGATGGTGCAGTGTGT 300	
DB	241 TCATTCCCGGGAGGATATACATGGAGAGGATCATGAACCTTTGAAGATGGTGCAGTGTGT 300	
QY	301 ACTGTACGATGATTCAGATCCAGGACATGTTTCATCTACCATGTCAAGTTCCT 360	
DB	301 ACTGTACGATGATTCAGATCCAGGACATGTTTCATCTACCATGTCAAGTTCCT 360	
QY	361 GGTTTGAACCTTCTCCCAAGTGACCTCTTATGAGAGAGAGACACAGGGCTGGGAACC 420	
DB	361 GGTTTGAACCTTCTCCCAAGTGACCTCTTATGAGAGAGAGACACAGGGCTGGGAACC 420	
QY	421 AACACTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTGTG 480	
DB	421 AACACTGAGCGTCTCTCTGACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTGTG 480	
QY	481 AAGTTAGAGAGGGTGGTCACTATTGTTGTAATTCAAATCTACTTACAGGCAAGGAG 540	
DB	481 AAGTTAGAGAGGGTGGTCACTATTGTTGTAATTCAAATCTACTTACAGGCAAGGAG 540	

QY	541 CCTGTGAAGATGCAGGGTATCACTATGTTGACCGCAAACTGGATGTAAACAATCAAC 600	
DB	541 CCTGTGAAGATGCAGGGTATCACTATGTTGACCGCAAACTGGATGTAAACAATCAAC 600	
QY	601 AAGGATTACACTTCCGTTGAGCAGCGTGAATAATTTCCATTGACGCAAACTTTGGTGGCC 660	
DB	601 AAGGATTACACTTCCGTTGAGCAGCGTGAATAATTTCCATTGACGCAAACTTTGGTGGCC 660	
RESULT 7		
ID	ABZ12069 standard; DNA; 660 BP.	
XX	AC ABZ12069;	
XX	22-JAN-2003 (first entry)	
DE	Colour Facilitating molecule (CFM) related sequence #SEQ ID 21.	
KW	Colour facilitating molecule; CFM; green fluorescent protein; GFP;	
KW	chromophore; biomatrix; transgenic animal; colouring agent;	
KW	flower industry; expression marker; reporter molecule; photon trap;	
XX	UV sink; sunsreen; ds.	
OS	Acropora aspera.	
XX	WO200270703-A2.	
XX	12-SEP-2002.	
XX	01-MAR-2002; 2002WO-GB000928.	
XX	02-MAR-2001; 2001US-0273227P.	
PR	21-MAR-2001; 2001AU-00003874.	
PR	15-OCT-2001; 2001US-0329816P.	
PA	(NUFA-) NUFARM LTD.	
PA	(UYOU) UNIV QUEENSLAND.	
PA	(JONE/) JONES E L.	
PI	Jones EL, Karan M, Brugliera F, Mason J, Dove SG;	
PI	Hoeigh-Guldberg IO, Prescott M;	
XX	WPI; 2002-740765/80.	
XX	Novel color-facilitating molecule for producing a biomatrix, has a	
PT	polypeptide which alone/along with molecules imparts altered visual	
PT	characteristics to cells in the absence of excitation by extraneous non-	
PT	white light.	
PS	Claim 6; Page 285-286; 510pp; English.	
XX	The invention relates to an isolated colour-facilitating molecule (CFM)	
CC	comprising a polypeptide which, in a cell, alone or together with one or	
CC	more other molecules imparts an altered visual characteristic to the cell	
CC	when visualised by a human eye in the absence of excitation by extraneous	
CC	non-white light or particle emission. CFMs are useful for producing a	
CC	transgenic animal which exhibits a novel colour e.g. sheep with blue or	
CC	red coloured fleece. They are useful for producing coloured plant	
CC	extracts, e.g. flavouring, beverage or juice or colouring agent. Other	
CC	uses include transducing or intensifying an image, providing additional	
CC	light for growing phototropic organisms e.g. algae and/or corals, for	
CC	coating materials that experience UV damage e.g. plastics and car	
CC	upholstery. CFMs are useful in the flower industry, in the development of	
CC	new varieties of flowering plants. Other contemplated uses include,	
CC	expression markers, general reporter molecules, photon traps, UV sinks or	
CC	in sunscreens. CFMs modify visible colour in edible and/or ornamental	
CC	funeral species, and in fruits and vegetables to enhance their	
CC	marketability. CFMs embedded in a gel matrix improve image quality in	
CC	situations of distorted light spectra (biomatrix). The first all-protein	
CC	chromophore to be isolated was Green Fluorescent protein (GFP). The	
CC	sequences given in records ABZ12068-ABZ12195 represent CFM related DNA	

```
CC sequences
XX SQ Sequence 660 BP; 193 A; 146 C; 163 G; 158 T; 0 U; 0 Other;
Query Match 77.1%; Score 648.8; DB 6; Length 660;
Best Local Similarity 98.9%; Pred. No. 1.5e-149;
Matches 653; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATGTCTAGGCGACGGTCAATGACAC 60
Db 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATGTCTAGGCGACGGTCAATGACAC 60
QY 61 TACTTTGAGGTCGAAGGCGATGGAAGAAAGAAACCTTACGAGGGGAGCAGACGGTAAAG 120
Db 61 TACTTTGAGGTCGAAGGCGATGGAAGAAAGAAACCTTACGAGGGGAGCAGACGGTAAAG 120
QY 121 CTGGCTGTCAACCAAGGGGAGCTCTGCCATTTGCTTGGGATATTTATCACCAGTGT 180
Db 121 CTGACTGTCAACCAAGGGGAGCTCTGCCATTTGCTTGGGATATTTATCACCAGTGT 180
QY 181 CAGTACGGAAGCATACCTTACCAAGTACCTCGAAGACATCCCTGACTATGTTAAAGCAG 240
Db 181 CAGTACGGAAGCATACCTTACCAAGTACCTCGAAGACATCCCTGACTATGTTAAAGCAG 240
QY 241 TCATTTCCCGGGAGATATACATGGGAGGATCATGAACCTTTGAAGATGGTCAAGTGT 300
Db 241 TCATTTCCCGGGAGGATATACATGGGAGGATCATGAACCTTTGAAGATGGTCAAGTGT 300
QY 301 ACTGTCAGCAATGANTCCAGATCCAGGCAACTGTTTCATCTACCATGTCAAGTCTCT 360
Db 301 ACTGTCAGCAATGANTCCAGGATCCAGGCAACTGTTTCATCTACCATGTCAAGTCTCT 360
QY 361 GGTTTGAACCTTCCCTCCAAATGGACCTGTATGACGAGAGGATCATGAACCTTTGAAGATGGTCAAGTCTCT 420
Db 361 GGTTTGAACCTTCCCTCCAAATGGACCTGTATGACGAGAGGATCATGAACCTTTGAAGATGGTCAAGTCTCT 420
QY 421 AACACTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGGAAACAACTTTATGCTCTG 480
Db 421 AACACTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGGAAACAACTTTATGCTCTG 480
QY 481 AAGTTAGAGAGGTTGGTCACTATTTGTTGTAATTCATCAATCTTCAAGCAAGGAAG 540
Db 481 AAGTTAGAGAGGTTGGTCACTATTTGTTGTAATTCATCAATCTTCAAGCAAGGAAG 540
QY 541 CCGTGAAGATGCCAGGGTATCAGTATGTCACCGCAACCTGGATGAACCAATCACAAC 600
Db 541 CCGTGAAGATGCCAGGGTATCAGTATGTCACCGCAACCTGGATGAACCAATCACAAC 600
QY 601 AAGGATTACACTTCGGTTGAGCAGCGTGAAATTTCCATTGACGCAAAACCTTTGTCGCC 660
Db 601 AAGGATTACACTTCGGTTGAGCAGCGTGAAATTTCCATTGACGCAAAACCTTTGTCGCC 660
RESULT 8
ABZ12156
ID ABZ12156 standard; DNA; 669 BP.
XX AC ABZ12156;
XX XX
XX XX
XX 22-JAN-2003 (first entry)
Dx Colour Facilitating molecule (CFM) related sequence #SEQ ID 211.
DE Colour facilitating molecule; CFM; green fluorescent protein; GFP;
DE KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
KW UV sink; sunsreen; ds.
XX Discosoma sp.
OS WO200270703-A2.
XX PN 12-SEP-2002.
XX PD
```

```
XX 01-MAR-2002; 2002WO-GB000928.
XX 02-MAR-2001; 2001US-0273227P.
PR 21-MAR-2001; 2001AU-00003874.
PR 15-OCT-2001; 2001US-0329816P.
XX (NUFA-) NUFARM LTD.
PA (UYOU) UNIV QUEENSLAND.
PA (JONE/) JONES E L.
XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
PI Hoegh-Guldberg IO, Prescott M;
XX WPI; 2002-740765/80.
XX Novel color-facilitating molecule for producing a biomatrix, has a
PT polypeptide which alone/along with molecules imparts altered visual
PT characteristics to cells in the absence of excitation by extraneous non-
PT white light.
XX Example 12; Page 482-483; 510pp; English.
XX The invention relates to an isolated colour-facilitating molecule (CFM)
CC comprising a polypeptide which, in a cell, alone or together with one or
CC more other molecules imparts an altered visual characteristic to the cell
CC when visualised by a human eye in the absence of excitation by extraneous
CC non-white light or particle emission. CFMs are useful for producing a
CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
CC red coloured fleece. They are useful for producing coloured plant
CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
CC uses include transducing or intensifying an image, providing additional
CC light for growing phototropic organisms e.g. algae and/or corals, for
CC coating materials that experience UV damage e.g. plastics and car
CC upholstery. CFMs are useful in the flower industry, in the development of
CC new varieties of flowering plants. Other contemplated uses include,
CC expression markers, general reporter molecules, photon traps, UV sinks or
CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
CC fungal species, and in fruits and vegetables to enhance their
CC marketability. CFMs embedded in a gel matrix improve image quality in
CC situations of distorted light spectra (biomatrix). The first all-protein
CC chromophore to be isolated was Green fluorescent protein (GFP). The
CC sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
CC sequences
XX SQ Sequence 669 BP; 195 A; 148 C; 167 G; 159 T; 0 U; 0 Other;
Query Match 77.1%; Score 648.8; DB 6; Length 669;
Best Local Similarity 98.9%; Pred. No. 1.5e-149;
Matches 653; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCTAGGCGACGGTCAATGACAC 60
Db 4 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCTAGGCGACGGTCAATGACAC 63
QY 61 TACTTTGAGGTCGAAGGCGATGGAAGAAAGAAAGCTTACGAGGGGAGCAGACGGTAAAG 120
Db 64 TACTTTGAGGTCGAAGGCGATGGAAGAAAGAAAGCTTACGAGGGGAGCAGACGGTAAAG 123
QY 121 CTGGCTGTCAACCAAGGGGAGCTCTGCCATTTGCTTGGGATATTTATCACCAGTGT 180
Db 124 CTGACTGTCAACCAAGGGGAGCTCTGCCATTTGCTTGGGATATTTATCACCAGTGT 183
QY 181 CAGTACGGAAGCATACCTTACCAAGTACCTCGAAGACATCCCTGACTATGTTAAAGCAG 240
Db 184 CAGTACGGAAGCATACCTTACCAAGTACCTCGAAGACATCCCTGACTATGTTAAAGCAG 243
QY 241 TCATTTCCCGGGAGATATACATGGGAGGATCATGAACCTTTGAAGATGGTCAAGTGT 300
Db 244 TCATTTCCCGGGAGGATATACATGGGAGGATCATGAACCTTTGAAGATGGTCAAGTGT 303
QY 301 ACTGTGAGCAATGATTCAGCATCCAAAGCACTGTTTCATCTACCATGTCAAGTCTCT 360
```

Db 304 ACTGTGAGCAATGATCCAGCATCCAGGCAACTGTTTCATCTACCATGTCAAGTTCTCT 363
 QY 361 GGTGTTGAACCTTCTCCCAATGACCTGTTATGACAGAAAGACACAGGGGTGGAAACCC 420
 Db 364 GGTGTTGAACCTTCTCCCAATGACCTGTTATGACAGAAAGACACAGGGGTGGAAACCC 423
 QY 421 AACACTGAGCGTCTCTTTGACAGAGATGGAATGCTGATAGAAACAATTTATGGCTCTG 480
 Db 424 AACACTGAGCGTCTCTTTGACAGAGATGGAATGCTGATAGAAACAATTTATGGCTCTG 483
 QY 481 AAGTTAGAGAGGTGCTACTATTGTTGTAATTTCAATCTACTTACAGGCAAGGAAG 540
 Db 484 AAGTTAGAGAGGTGCTACTATTGTTGTAATTTCAATCTACTTACAGGCAAGGAAG 543
 QY 541 CCTGTGAAGATGCCAGGCTATCACTATGTTGACCGCAAACTGGATGTAACCAATCAACAAC 600
 Db 544 CCTGTGAAGATGCCAGGCTATCACTATGTTGACCGCAAACTGGATGTAACCAATCAACAAC 603
 QY 601 AAGGATTACACTTCCGTTGAGCAGCGTGAATTTCCATTGACGCAAACTTTGTCGCC 660
 Db 604 AAGGATTACACTTCCGTTGAGCAGCGTGAATTTCCATTGACGCAAACTTTGTCGCC 663

RESULT 9

ABZ12070
 ID ABZ12070 standard; DNA; 663 BP.

AC ABZ12070;

XX 22-JAN-2003 (first entry)

DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 23.

XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunscreen; ds.

XX Acropora aspera.

XX WO200270703-A2.

XX 12-SEP-2002.

XX 01-MAR-2002; 2002WO-GB000928.

XX 02-MAR-2001; 2001US-0273227P.

XX 21-MAR-2001; 2001AU-00003874.

XX 15-OCT-2001; 2001US-0329816P.

XX (NUFA-) NUFARM LTD.

XX (UYQU) UNIV QUEENSLAND.

XX (JONE/) JONES E L.

XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;

XX Hoegh-Guldberg IO, Prescott M;

XX WPI; 2002-740765/80.

XX Claim 6; Page 287-288; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)
 CC comprising a polypeptide which, in a cell, alone or together with one or
 CC more other molecules imparts an altered visual characteristic to the cell
 CC when visualised by a human eye in the absence of excitation by extraneous
 CC non-white light or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant

CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development of
 CC new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
 CC sequences

XX Sequence 663 BP; 193 A; 145 C; 164 G; 161 T; 0 U; 0 Other;

Query Match 77.0%; Score 647.6; DB 6; Length 663;

Best Local Similarity 98.6%; Pred. No. 2.9e-149;

Matches 653; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 TCCGTTATCGTAAACAGATGACCTACAAAGTTTATATGTGACGACACGTCATGGACAC 60
 Db 1 TCCGTTATCGTAAACAGATGACCTACAAAGTTTATATGTGACGACACGTCATGGACAC 60
 QY 61 TACTTTGAGTTCGAAGCGATGGAAGAAAGAAAGCCTTACGAGGGGAGCAGCGTAAAG 120
 Db 61 TACTTTGAGTTCGAAGCGATGGAAGAAAGAAAGCCTTACGAGGGGAGCAGCGTAAAG 120
 QY 121 CTGGCTGTCCCAAGGGGGACCTCTGCCATTTCTCTGGGATATTTATATACCCACAGTGT 180
 Db 121 CTGACTGTCCCAAGGGGGACCTCTGCCATTTCTCTGGGATATTTATATACCCACAGTGT 180
 QY 181 CAGTACGGAAGCATACCATTCACCAAGTACCTTGAAGACATCCCTGACTATGTAAGCAG 240
 Db 181 CAGTACGGAAGCATACCATTCACCAAGTACCTTGAAGACATCCCTGACTATGTAAGCAG 240
 QY 241 TCATTCCCGGGAGATATACATGGGAGGAGATCATGAACCTTTGAAGATGGTGCACTGTGT 300
 Db 241 TCATTCCCGGGAGGATATACATGGGAGGAGATCATGAACCTTTGAAGATGGTGCACTGTGT 300
 QY 301 ACTGTGACGATGATTCAGCATCCAGGCACTGTTTTCATCTACCATGTCAAGTTCTCT 360
 Db 301 ACTGTGACGATGATTCAGCATCCAGGCACTGTTTTCATCTACCATGTCAAGTTCTCT 360
 QY 361 GGTGTTGAACCTTCTCCCAATGGACCTGTATGACAGAAAGACACAGGGCTGGAAACCC 420
 Db 361 GGTGTTGAACCTTCTCCCAATGGACCTGTATGACAGAAAGACACAGGGCTGGAAACCC 420
 QY 421 AACACTGAGCGTCTCTTTGACAGAGATGGAATGCTGATAGGAAACAACCTTTATGGCTCTG 480
 Db 421 AACACTGAGCGTCTCTTTGACAGAGATGGAATGCTGATAGGAAACAACCTTTATGGCTCTG 480
 QY 481 AAGTTAGAGAGGTGGTCACTATTGTTGTAATTTCAATCTACTTTACAAGGCAAGGAAG 540
 Db 481 AAGTTAGAGAGGTGGTCACTATTGTTGTAATTTCAATCTACTTTACAAGGCAAGGAAG 540
 QY 541 CCTGTGAAGATGCCAGGCTATCACTATGTTGACCGCAAACTGGATGTAACCAATCAACAAC 600
 Db 541 CCTGTGAAGATGCCAGGCTATCACTATGTTGACCGCAAACTGGATGTAACCAATCAACAAC 600
 QY 601 AAGGATTACACTTCCGTTGAGCAGCGTGAATTTCCATTGACGCAAACTTTGTCGCC 660
 Db 601 AAGGATTACACTTCCGTTGAGCAGCGTGAATTTCCATTGACGCAAACTTTGTCGCC 660
 QY 661 TG 662
 Db 661 TG 662

RESULT 10

ABZ12072
 ID ABZ12072 standard; DNA; 660 BP.

XX AC ABZ12072;
 XX DT 07-AUG-2003 (revised)
 XX DT 22-JAN-2003 (first entry)
 XX DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 27.
 XX DE Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunscreen; ds.
 XX ACanthastrea sp.
 XX OS WO200270703-A2.
 XX PN 12-SEP-2002.
 XX PD 01-MAR-2002; 2002WO-GB000928.
 XX PF 02-MAR-2001; 2001US-0273227P.
 XX PR 21-MAR-2001; 2001AU-00003874.
 XX ER 15-OCT-2001; 2001US-0329816P.
 XX (NUFA-) NUFARM LTD.
 XX PA (UYQU) UNIV QUEENSLAND.
 XX PA (JONE/) JONES E L.
 XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
 FI Hoegh-Guldberg IO, Prescott M;
 PI WPI; 2002-740765/80.
 XX Novel color-facilitating molecule for producing a biomatrix, has a
 PT polypeptide which alone/along with molecules imparts altered visual
 PT characteristics to cells in the absence of excitation by extraneous non-
 PT white light.
 XX Claim 6; Page 292-293; 51opp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)
 CC comprising a polypeptide which, in a cell, alone or together with one or
 CC more other molecules imparts an altered visual characteristic to the cell
 CC when visualised by a human eye in the absence of excitation by extraneous
 CC non-white light or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant
 CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development of
 CC new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
 CC sequences. (Updated on 07-AUG-2003 to correct OS field.)

XX Sequence 660 BP; 191 A; 147 C; 164 G; 158 T; 0 U; 0 Other;
 Query Match 77.0%; Score 647.2; DB 6; Length 660;
 Best Local Similarity 98.8%; Pred. No. 3.6e-149;
 Matches 652; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 TCCGTTATCGCTAACACAGATGACCTACAAAGTTTATGTCAGGACGGTCAATGGACAC 60
 DB 1 TCCGTTATCGCTAACACAGATGACCTACAAAGTTTATGTCAGGACGGTCAATGGACAC 60

QY 61 TACTTTGAGGTGCGAAGCGGATGGAAGAAAGCCCTTACGAGGGGAGCAGCGGTAAGG 120
 DB |||||
 QY 61 TACTTTGAGGTGCGAAGCGGATGGAAGAAAGCCCTTACGAGGGGAGCAGCGGTAAGG 120
 DB |||||
 QY 121 CTGGCTGTCAACAAAGGCGGACCTCTGCCATTTCCTTGGGATATTTTATCACCACAGTGT 180
 DB |||||
 QY 121 CTGACTGTCAACAAAGGCGGACCTCTGCCATTTCCTTGGGATATTTTATCACCACAGTGT 180
 DB |||||
 QY 181 CAGTACGGAAGCATACCAATTCACCAAGTACCTCGCTGAAGACATCCCTGATATGTAAGCAG 240
 DB |||||
 QY 181 CAGTACGGAAGCATACCAATTCACCAAGTACCTCGCTGAAGACATCCCTGATATGTAAGCAG 240
 DB |||||
 QY 241 TCATTCCCGGGAGATATACATGGGAGGATCATGAACCTTTGAGAGTGTGTCAGTGTGT 300
 DB |||||
 QY 241 TCATTCCCGGGAGATATACATGGGAGGATCATGAACCTTTGAGAGTGTGTCAGTGTGT 300
 DB |||||
 QY 301 ACTGTTCAGCAATGATTCAGCATCCAGGCAACTGTTTTCATCTACCATGTCAAAGTTCTCT 360
 DB |||||
 QY 301 ACTGTTCAGCAATGATTCAGCATCCAGGCAACTGTTTTCATCTACCATGTCAAAGTTCTCT 360
 DB |||||
 QY 361 GGTGTGAACCTTCTCCCAATGGACCTGTTATGAGAGAAAGACACAGGGCTGGGAACC 420
 DB |||||
 QY 361 GGTGTGAACCTTCTCCCAATGGACCTGTTATGAGAGAAAGACACAGGGCTGGGAACC 420
 DB |||||
 QY 421 AACACTGAGCGTCTCTTTGACAGAGATGGAATGCTGATAGGAACAACCTTTATGCTCTG 480
 DB |||||
 QY 421 AACACTGAGCGTCTCTTTGACAGAGATGGAATGCTGATAGGAACAACCTTTATGCTCTG 480
 DB |||||
 QY 481 AAGTTAGAGAGGTGTCACATTTTGTGTAATTCAAATCTACTTACAGGCAAGGAAG 540
 DB |||||
 QY 481 AAGTTAGAGAGGTGTCACATTTTGTGTAATTCAAATCTACTTACAGGCAAGGAAG 540
 DB |||||
 QY 541 CCTGTGAAGATGCCAGGATATCACTATGTTGACCGCAAACTGGATGTAAACAATCACAAC 600
 DB |||||
 QY 541 CCTGTGAAGATGCCAGGATATCACTATGTTGACCGCAAACTGGATGTAAACAATCACAAC 600
 DB |||||
 QY 601 AAGGATTACACTTCGTTGACAGCGTGAAATTTCCATTGACCGCAAACTTTGTCGCC 660
 DB |||||
 QY 601 AAGGATTACACTTCGTTGACAGCGTGAAATTTCCATTGACCGCAAACTTTGTCGCC 660
 DB |||||

RESULT 11
 ABZ12083
 ID ABZ12083 standard; DNA; 660 BP.
 XX
 AC ABZ12083;
 XX
 DT 22-JAN-2003 (first entry)
 XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 49.
 DE Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 DE chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunscreen; ds.
 XX Millepora sp.
 XX OS WO200270703-A2.
 XX PN 12-SEP-2002.
 XX PF 01-MAR-2002; 2002WO-GB000928.
 XX PR 02-MAR-2001; 2001US-0273227P.
 XX PR 21-MAR-2001; 2001AU-00003874.
 XX PR 15-OCT-2001; 2001US-0329816P.
 XX (NUFA-) NUFARM LTD.
 XX PA (UYQU) UNIV QUEENSLAND.
 XX PA (JONE/) JONES E L.
 XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;

PI	Hoegh-Guldberg IO, Prescott M;	481	AAGTTAGAGGAGGTGGTCACTATTATGTGAATTCAAATCTTACTTACAAAGCGAAGGAG	540
XX	WPI; 2002-740765/80.	541	CCTGTGAAGATGCCAGGGTATCACTATGTTGACGCAAACTGGATGTAAACCAATCAACAAC	600
XX	Novel color-facilitating molecule for producing a biomatrix, has a	541		600
PT	polypeptide which alone/along with molecules imparts altered visual	601	CCTGTGAAGATGCCAGGGTATCACTATGTTGACGCAAACTGGATGTAAACCAATCAACAAC	660
PT	characteristics to cells in the absence of excitation by extraneous non-	601		660
PT	white light.	601	AAGGATTACACTTCCGTTGAGCAGCGTGAATTTCCATTGACGCAAACTGTTGGTGGCC	660
XX	Claim 6; Page 324-325; 510pp; English.	601		660
XX	The invention relates to an isolated colour-facilitating molecule (CFM)	RESULT 12		
CC	comprising a polypeptide which, in a cell, alone or together with one or	ABZ12098		
CC	more other molecules imparts an altered visual characteristic to the cell	ID	ABZ12098 standard; DNA; 660 BP.	
CC	when visualised by a human eye in the absence of excitation by extraneous	XX	XX	
CC	non-white light or particle emission. CFMs are useful for producing a	AC	ABZ12098;	
CC	transgenic animal which exhibits a novel colour e.g. sheep with blue or	XX	07-AUG-2003 (revised)	
CC	red coloured fleece. They are useful for producing coloured plant	DT	22-JAN-2003 (first entry)	
CC	extracts, e.g. flavouring, beverage or juice or colouring agent. Other	XX	Colour Facilitating molecule (CFM) related sequence #SEQ ID 81.	
CC	uses include transducing or intensifying an image, providing additional	XX	Colour facilitating molecule: CFM; green fluorescent protein; GFP;	
CC	light for growing phototropic organisms e.g. algae and/or corals, for	XX	chromophore; biomatrix; transgenic animal; colouring agent;	
CC	coating materials that experience UV damage e.g. plastics and car	XX	flower industry; expression marker; reporter molecule; photon trap;	
CC	upholstery. CFMs are useful in the flower industry, in the development of	XX	UV sink; sunscreen; ds.	
CC	new varieties of flowering plants. Other contemplated uses include;	OS	Pavona decussata.	
CC	expression markers, general reporter molecules, photon traps, UV sinks or	XX	WO200270703-A2.	
CC	in sunscreens. CFMs modify visible colour in edible and/or ornamental	XX	12-SEP-2002.	
CC	fungi species, and in fruits and vegetables to enhance their	XX	01-MAR-2002; 2002WO-GB000928.	
CC	marketability. CFMs embedded in a gel matrix improve image quality in	XX	02-MAR-2001; 2001US-0273227P.	
CC	situations of distorted light spectra (biomatrix). The first all-protein	XX	21-MAR-2001; 2001AU-00003874.	
CC	chromophore to be isolated was Green Fluorescent protein (GFP). The	XX	15-OCT-2001; 2001US-039816P.	
CC	sequences given in records ABZ12068-ABZ12195 represent CFM related DNA	XX	(NUFA-) NUFARM LTD.	
XX	sequences	XX	(UYOU) UNIV QUEENSLAND.	
XX	Sequence 660 BP; 193 A; 146 C; 161 G; 160 T; 0 U; 0 Other;	XX	(JONE/) JONES E L.	
SQ		PI	Jones EL, Karan M, Brugliera F, Mason J, Dove SG;	
	Query Match 76.2%; Score 640.8; DB 6; Length 660;	PI	Hoegh-Guldberg IO, Prescott M;	
	Best Local Similarity 98.2%; Pred. No. 1.4e-147;	XX	WPI; 2002-740765/80.	
	Matches 648; Conservative 0; Mismatches 12; Indels 0; Gaps 0;	XX	Novel color-facilitating molecule for producing a biomatrix, has a	
QY	1 TCCGTTATCGTAAACAGATGACCTACAAAGCTTATATGTCAGGCGAGCTCAATGGACAC	XX	polypeptide which alone/along with molecules imparts altered visual	
DB	1 TCCGTTATCGTAAACAGATGACCTACAAAGCTTATATGTCAGGCGAGCTCAATGGACAC	XX	characteristics to cells in the absence of excitation by extraneous non-	
QY	61 TACTTTGAGTTCGAGGCGATGGAAGAAAGCTTACGAGGGGGAGCAGACGCTAAGG	XX	white light.	
DB	61 TACTTTGAGTTCGAGGCGATGGAAGAAAGCTTACGAGGGGGAGCAGACGCTAAGG	XX	Claim 6; Page 360-361; 510pp; English.	
QY	121 CTGCTGTCCACAGGGGGACCTCTGCCATTTGCTGGGATATTTATCACACAGTGT	XX	The invention relates to an isolated colour-facilitating molecule (CFM)	
DB	121 CTGCTGTCCACAGGGGGACCTCTGCCATTTGCTGGGATATTTATCACACAGTGT	XX	comprising a polypeptide which, in a cell, alone or together with one or	
QY	181 CAGTACGGAAGCATACCAAGTACCAAGTACCAAGTACCAAGTACCAAGTACCAAGTAC	XX	more other molecules imparts an altered visual characteristic to the cell	
DB	181 CAGTACGGAAGCATACCAAGTACCAAGTACCAAGTACCAAGTACCAAGTACCAAGTAC	XX	when visualised by a human eye in the absence of excitation by extraneous	
QY	241 TCATTCCCGGGAGATATACATGGAGGAGGATCATGAATTTGAAGATGTTGAGTGTGT	XX	non-white light or particle emission. CFMs are useful for producing a	
DB	241 TCATTCCCGGGAGATATACATGGAGGAGGATCATGAATTTGAAGATGTTGAGTGTGT	XX	transgenic animal which exhibits a novel colour e.g. sheep with blue or	
QY	301 ACTGTGACATGATTCAGCATCCAGGCACTGTTTCATCTACCATGTCAGTTCTCT	XX	red coloured fleece. They are useful for producing coloured plant	
DB	301 ACTGTGACATGATTCAGCATCCAGGCACTGTTTCATCTACCATGTCAGTTCTCT	XX	extracts, e.g. flavouring, beverage or juice or colouring agent. Other	
QY	361 GGTTTGAATTTCTCTCCCAATGGACCTGTTATGAGAGAGAGACACAGGGCTGGGAACC	XX	uses include transducing or intensifying an image, providing additional	
DB	361 GGTTTGAATTTCTCTCCCAATGGACCTGTTATGAGAGAGAGACACAGGGCTGGGAACC	XX	light for growing phototropic organisms e.g. algae and/or corals, for	
QY	421 AACCTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGGAACAACTTATGCTCTG	XX	coating materials that experience UV damage e.g. plastics and car	
DB	421 AACCTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGGAACAACTTATGCTCTG	XX	upholstery. CFMs are useful in the flower industry, in the development of	
QY	481 AAGTTAGAGGAGGTGGTCACTATTGTGTAATTCAAATCTTACTTACAAAGCGAAGGAG	XX	new varieties of flowering plants. Other contemplated uses include;	
		XX	expression markers, general reporter molecules, photon traps, UV sinks or	
		XX	in sunscreens. CFMs modify visible colour in edible and/or ornamental	
		XX	fungi species, and in fruits and vegetables to enhance their	
		XX	marketability. CFMs embedded in a gel matrix improve image quality in	
		XX	situations of distorted light spectra (biomatrix). The first all-protein	
		XX	chromophore to be isolated was Green Fluorescent protein (GFP). The	
		XX	sequences given in records ABZ12068-ABZ12195 represent CFM related DNA	
		XX	sequences	

CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
 CC sequences. (Updated on 07-AUG-2003 to correct OS field.)
 XX

SQ Sequence 660 BP; 191 A; 147 C; 164 G; 158 T; 0 U; 0 Other;
 Query Match 75.8%; Score 637.6; DB 6; Length 660;
 Best Local Similarity 97.9%; Pred. No. 8.3e-147;
 Matches 646; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
 QY 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGACGGTCAATGGACAC 60
 Db 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGACGGTCAATGGACAC 60
 QY 61 TACTTTGAGTGCAGAGCGATGAAAGAAAGAAAGCCCTTACGAGGGGAGCAGCGTAAAG 120
 Db 61 TACTTTGAGTGCAGAGCGATGAAAGAAAGAAAGCCCTTACGAGGGGAGCAGCGTAAAG 120
 QY 121 CTGGCTGTCAACAGGGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGT 180
 Db 121 CTCAGTGTCAACAGGGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGT 180
 QY 181 CAGTACGAGGATACCAATTCACCAAGTACCTGAGAGATCCCTGACTATGTAAGCAG 240
 Db 181 CAGTACGAGGATACCAATTCACCAAGTACCTGAGAGATCCCTGACTATGTAAGCAG 240
 QY 241 TCATTCCCGGGAGATATACATGGGAGGATCATGAACCTTTGAAGATGGTCAAGTGT 300
 Db 241 TCATTCCCGGGAGATATACATGGGAGGATCATGAACCTTTGAAGATGGTCAAGTGT 300
 QY 301 ACTGTCAAGATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAGTTCCT 360
 Db 301 ACTGTCAAGATGATTCAGCATCCAGGCAACTGTTTCATCTACCATGTCAGTTCCT 360
 QY 361 GGTGTGAACCTTCTCCCAAGTGCATGAGAGGATCATGAACCTTTGAAGATGGTCAAGTGT 420
 Db 361 GGTGTGAACCTTCTCCCAAGTGCATGAGAGGATCATGAACCTTTGAAGATGGTCAAGTGT 420
 QY 421 AACACTGAGCGTCTCTTTGACAGGATGGAATGCTGATAGGAACAACTTTATGCTCTG 480
 Db 421 AACACTGAGCGTCTCTTTGACAGGATGGAATGCTGATAGGAACAACTTTATGCTCTG 480
 QY 481 AAGTTAGAGAGGTGGTCACTATTGTGTGAATTCAAATCTCTTCAAGGCAAGGAAG 540
 Db 481 AAGTTAGAGAGGTGGTCACTATTGTGTGAATTCAAATCTCTTCAAGGCAAGGAAG 540
 QY 541 CCTGTGAGATGCCAGGTATCACTATGTTGACCGCAAACTGGATGAACCAATCAAC 600
 Db 541 CCTGTGAGATGCCAGGTATCACTATGTTGACCGCAAACTGGATGAACCAATCAAC 600
 QY 601 AAGGATTACACTTCCGTTGACAGCGTGAATTTCCATTGACGCAAACTTTGCTGCC 660
 Db 601 AAGGATTACACTTCCGTTGACAGCGTGAATTTCCATTGACGCAAACTTTGCTGCC 660

RESULT 13
 ABZ12155

ID ABZ12155 standard; DNA; 669 BP.

AC ABZ12155;

XX 22-JAN-2003 (first entry)

Colour Facilitating molecule (CFM) related sequence #SEQ ID 209.

Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 chromophore; biomatrix; transgenic animal; colouring agent;
 flower industry; expression marker; reporter molecule; photon trap;
 UV sink; sunsreen; ds.

OS Acropora sp.

XX WO200270703-A2.

XX 12-SEP-2002.

XX 01-MAR-2002; 2002WO-GB000928.

XX 02-MAR-2001; 2001US-0273227P.

XX 21-MAR-2001; 2001AU-00003874.

XX 15-OCT-2001; 2001US-0329816P.

XX (NUFA-) NUFARM LTD.

XX (UYQU) UNIV QUEENSLAND.

XX (JONE/) JONES E L.

PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;

PI Hoegn-Guldberg IO, Prescott M;

XX WPI; 2002-740765/80.

Novel color-facilitating molecule for producing a biomatrix, has a
 polypeptide which alone/along with molecules imparts altered visual
 characteristics to cells in the absence of excitation by extraneous non-
 white light.

Example 11; Page 481; 510pp; English.

The invention relates to an isolated colour-facilitating molecule (CFM)
 comprising a polypeptide which, in a cell, alone or together with one or
 more other molecules imparts an altered visual characteristic to the cell
 when visualised by a human eye in the absence of excitation by extraneous
 non-white light or particle emission. CFMs are useful for producing a
 transgenic animal which exhibits a novel colour e.g. sheep with blue or
 red coloured fleeces. They are useful for producing coloured plant
 extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 uses include transducing or intensifying an image, providing additional
 light for growing phototropic organisms e.g. algae and/or corals, for
 coating materials that experience UV damage e.g. plastics and car
 upholstery. CFMs are useful in the flower industry, in the development of
 new varieties of flowering plants. Other contemplated uses include,
 expression markers, general reporter molecules, photon traps, UV sinks or
 in sunscreens. CFMs modify visible colour in edible and/or ornamental
 fungal species, and in fruits and vegetables to enhance their
 marketability. CFMs embedded in a gel matrix improve image quality in
 situations of distorted light spectra (biomatrix). The first all-protein
 chromophore to be isolated was Green Fluorescent protein (GFP). The
 sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
 sequences

Sequence 669 BP; 195 A; 148 C; 166 G; 160 T; 0 U; 0 Other;

Query Match 75.8%; Score 637.6; DB 6; Length 669;

Best Local Similarity 97.9%; Pred. No. 8.4e-147;

Matches 646; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGACGGTCAATGGACAC 60

Db 4 TCCGTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGACGGTCAATGGACAC 63

QY 61 TACTTTGAGTGCAGAGCGATGAAAGAAAGCCCTTACGAGGGGAGCAGCGTAAAG 120

Db 64 TACTTTGAGTGCAGAGCGATGAAAGAAAGCCCTTACGAGGGGAGCAGCGTAAAG 123

QY 121 CTGGCTGTCAACAGGGGACCTCTGCCATTTGCTTGGGATATTTTATACACCAAGTGT 180

Db 124 CTCAGTGTCAACAGGGGACCTCTGCCATTTGCTTGGGATATTTTATACACCAAGTGT 183

QY 181 CAGTACGAGGATACCAATTCACCAAGTACCTGAAGACATCCCTGACTATGTAAGCAG 240

Db 184 CAGTACGAGGATACCAATTCACCAAGTACCTGAAGACATCCCTGACTATGTAAGCAG 243

QY 241 TCATTCCCGGGAGATATACATGGGAGGATCATGAACCTTTGAAGATGGTCAAGTGT 300

Db 244 TCATTCCCGGGAGATATACATGGGAGGATCATGAACCTTTGAAGATGGTCAAGTGT 303

301 ACTGTGACGAATGATCCAGCATCCAGGCAACTGTTTTCATCTACCATGTCAAGTTCTCT 360
304 ACTGTGACGAATGATCCAGCATCCAGGCAACTGTTTTCATCTACCATGTCAAGTTCTCT 363
361 GGTGTGAACCTTCTCCCAATGACCTGTTATGTCAGAAAGACACAGGGCTGGGAACCC 420
364 GGTGTGAACCTTCTCCCAATGACCTGTTATGTCAGAAAGACACAGGGCTGGGAACCC 423
421 AACACTGACGCTCTCTTTCACGAGATGGAATGCTGATAGGAACAACTTTATGGCTCTG 480
424 AACACTGACGCTCTCTTTCACGAGATGGAATGCTGATAGGAACAACTTTATGGCTCTG 483
481 AAGTTAGAAGGAGGTGCTCACTATTGTGTGAATTCAAATCTACTTCAAGGCAAGGAAG 540
484 AAGTTAGAAGGAGGTGCTCACTATTGTGTGAATTCAAATCTACTTCAAGGCAAGGAAG 543
541 CCTGTGAAGATGCCAGGGTATCACTATGTTGACCGGAAACTGATGTAACCAATCAACAAC 600
544 CCTGTGAAGATGCCAGGGTATCACTATGTTGACCGGAAACTGATGTAACCAATCAACAAC 603
601 AAGGATTACACTTCCGTTGACGAGCGTGAAATTTCCATTCACGCAAACTTTGTCGCC 660
604 AAGGATTACACTTCCGTTGACGAGCGTGAAATTTCCATTCACGCAAACTTTGTCGCC 663

RESULT 14
ABZ12158
ID ABZ12158 standard; DNA; 669 BP.
XX AC ABZ12158;
XX DT 22-JAN-2003 (first entry)
XX DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 221.
XX KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
XX KW chromophore; biomatrix; transgenic animal; colouring agent;
XX KW flower industry; expression marker; reporter molecule; photon trap;
XX KW UV sink; sunscreen; ds.
XX OS Tubastrea sp.
XX FN WO200270703-A2.
XX PD 12-SEP-2002.
XX PF 01-MAR-2002; 2002WO-GB000928.
XX PR 02-MAR-2001; 2001US-0273227P.
XX PR 21-MAR-2001; 2001AU-00003874.
XX PR 15-OCT-2001; 2001US-0329816P.
XX PA (NUFA-) NUFARM LTD.
XX PA (UQU) UNIV QUEENSLAND.
XX PA (JONE)/ JONES E L.
XX PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
XX PI Hoegh-Guldberg IO, Prescott M;
XX WPI; 2002-740765/80.
XX DR Novel color-facilitating molecule for producing a biomatrix, has a
XX PT polypeptide which alone/along with molecules imparts altered visual
XX PT characteristics to cells in the absence of excitation by extraneous non-
XX PT white light.
XX PS Example 18; Page 490; 510pp; English.
XX CC The invention relates to an isolated colour-facilitating molecule (CFM)
XX CC comprising a polypeptide which, in a cell, alone or together with one or
XX CC more other molecules imparts an altered visual characteristic to the cell
XX CC when visualised by a human eye in the absence of excitation by extraneous
XX CC non-white light or particle emission. CFMs are useful for producing a

transgenic animal which exhibits a novel colour e.g. sheep with blue or
red coloured fleece. They are useful for producing coloured plant
extracts, e.g. flavouring, beverage or juice or colouring agent. Other
uses include transducing or intensifying an image, providing additional
light for growing phototropic organisms e.g. algae and/or corals, for
coating materials that experience UV damage e.g. plastics and car
upholstery. CFMs are useful in the flower industry, in the development of
new varieties of flowering plants. Other contemplated uses include,
expression markers, general reporter molecules, photon traps, UV sinks or
in sunscreens. CFMs modify visible colour in edible and/or ornamental
fungal species, and in fruits and vegetables to enhance their
marketability. CFMs embedded in a gel matrix improve image quality in
situations of distorted light spectra (biomatrix). The first all-protein
chromophore to be isolated was Green Fluorescent protein (GFP). The
sequences given in records ABZ12068-ABZ12195 represent CFM related DNA
sequences
SQ Sequence 669 BP; 189 A; 147 C; 170 G; 163 T; 0 U; 0 Other;
Query Match 75.6%; Score 636; DB 6; Length 669;
Best Local Similarity 97.7%; Pred. No. 2.1e-146;
Matches 645; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 1 TCCGTTATCGTAAACAGATGACCTACAAAGTTTATATGTCAGGACGGTCAATGGACAC 60
Db 4 TCCGTTATCGTAAACAGATGACCTACAAAGTTTATATGTCAGGACGGTCAATGGACAC 63
Qy 61 TACTTTCAGGTTCGAAGCGGATGAAAAGAAAGCCTTACGAGGGGAGCAGACGGTAAGG 120
Db 64 TACTTTCAGGTTCGAAGCGGATGAAAAGAAAGCCTTACGAGGGGAGCAGACGGTAAGG 123
Qy 121 CTGGCTGTCAACAAAGGGGACCTCTGCCATTTCTTGGGATATTTATACCAACAGTGT 180
Db 124 CTGGCTGTCAACAAAGGGGACCTCTGCCATTTCTTGGGATATTTATACCAACAGTGT 183
Qy 181 CAGTACCGAAGCATACCATTCACCAAGTACCTCCTGAAGACATCCCTGACTATGTAAGCAG 240
Db 184 CAGTACCGAAGCATACCATTCACCAAGTACCTCCTGAAGACATCCCTGACTATGTAAGCAG 243
Qy 241 TCATTCCCGGGGAGATATACATGGGAGAGGATCATGAACCTTTGAAGATGGTGCAGTGTGT 300
Db 244 TCATTCCCGGGGAGATATACATGGGAGAGGATCATGAACCTTTGAAGATGGTGCAGTGTGT 303
Qy 301 ACTGTGACGAATGATCCAGCATCCAGGCAACTGTTTTCATCTACCATGTCAAGTTCTCT 360
Db 304 ACTGTGACGAATGATCCAGCATCCAGGCAACTGTTTTCATCTACCATGTCAAGTTCTCT 363
Qy 361 GGTGTGAACCTTCTCCCAATGACCTGTTATGTCAGAAAGACACAGGGCTGGGAACCC 420
Db 364 GGTGTGAACCTTCTCCCAATGACCTGTTATGTCAGAAAGACACAGGGCTGGGAACCC 423
Qy 421 AACACTGACGCTCTCTTTCACGAGATGGAATGCTGATAGGAACAACTTTATGGCTCTG 480
Db 424 AACACTGACGCTCTCTTTCACGAGATGGAATGCTGATAGGAACAACTTTATGGCTCTG 483
Qy 481 AAGTTAGAAGGAGGTGCTCACTATTGTGTGAATTCAAATCTACTTCAAGGCAAGGAAG 540
Db 484 AAGTTAGAAGGAGGTGCTCACTATTGTGTGAATTCAAATCTACTTCAAGGCAAGGAAG 543
Qy 541 CCTGTGAAGATGCCAGGGTATCACTATGTTGACCGGAAACTGATGTAACCAATCAACAAC 600
Db 544 CCTGTGAAGATGCCAGGGTATCACTATGTTGACCGGAAACTGATGTAACCAATCAACAAC 603
Qy 601 AAGGATTACACTTCCGTTGACGAGCGTGAAATTTCCATTCACGCAAACTTTGTCGCC 660
Db 604 AAGGATTACACTTCCGTTGACGAGCGTGAAATTTCCATTCACGCAAACTTTGTCGCC 663
RESULT 15
ABZ12158
ID ABZ12158 standard; DNA; 669 BP.
XX AC ABZ12158;
XX DT 22-JAN-2003 (first entry)
XX DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 221.
XX KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
XX KW chromophore; biomatrix; transgenic animal; colouring agent;
XX KW flower industry; expression marker; reporter molecule; photon trap;
XX KW UV sink; sunscreen; ds.
XX OS Tubastrea sp.
XX FN WO200270703-A2.
XX PD 12-SEP-2002.
XX PF 01-MAR-2002; 2002WO-GB000928.
XX PR 02-MAR-2001; 2001US-0273227P.
XX PR 21-MAR-2001; 2001AU-00003874.
XX PR 15-OCT-2001; 2001US-0329816P.
XX PA (NUFA-) NUFARM LTD.
XX PA (UQU) UNIV QUEENSLAND.
XX PA (JONE)/ JONES E L.
XX PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
XX PI Hoegh-Guldberg IO, Prescott M;
XX WPI; 2002-740765/80.
XX DR Novel color-facilitating molecule for producing a biomatrix, has a
XX PT polypeptide which alone/along with molecules imparts altered visual
XX PT characteristics to cells in the absence of excitation by extraneous non-
XX PT white light.
XX PS Example 18; Page 490; 510pp; English.
XX CC The invention relates to an isolated colour-facilitating molecule (CFM)
XX CC comprising a polypeptide which, in a cell, alone or together with one or
XX CC more other molecules imparts an altered visual characteristic to the cell
XX CC when visualised by a human eye in the absence of excitation by extraneous
XX CC non-white light or particle emission. CFMs are useful for producing a

XX 22-JAN-2003 (first entry)
 XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 215.
 DE
 XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 XX chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunscreen; ds.
 XX
 XX Tubastrea sp.
 OS
 XX WO200270703-A2.
 XX
 XX 12-SEP-2002.
 PD
 XX
 XX 01-MAR-2002; 2002WO-GB000928.
 EF
 XX 02-MAR-2001; 2001US-0273227P.
 XX
 XX 21-MAR-2001; 2001AU-00003874.
 PR
 XX 15-OCT-2001; 2001US-0329816P.
 PR
 XX (NUFA-) NUFARM LTD.
 PA
 PA (UYQU) UNIV QUEENSLAND.
 PA (JONE/) JONES E L.
 XX
 XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
 PI Hoegh-Guldberg IO, Prescott M;
 PI
 XX MPI; 2002-740765/80.
 DR
 XX Novel color-facilitating molecule for producing a biomatrix, has a
 XX polypeptide which alone/along with molecules imparts altered visual
 PT characteristics to cells in the absence of excitation by extraneous non-
 PT white light.
 PT
 XX Example 18; Page 485-486; 510pp; English.
 PS
 XX The invention relates to an isolated colour-facilitating molecule (CFM)
 CC comprising a polypeptide which, in a cell, alone or together with one or
 CC more other molecules imparts an altered visual characteristic to the cell
 CC when visualised by a human eye in the absence of excitation by extraneous
 CC non-white light or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant
 CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development of
 CC new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC sequences given in records AB212068-AB212195 represent CFM related DNA
 CC sequences
 XX
 XX Sequence 669 BP; 194 A; 148 C; 167 G; 160 T; 0 U; 0 Other;
 SQ
 Query Match 75.6%; Score 636; DB 6; Length 669;
 Best Local Similarity 97.7%; Pred. No. 2.1e-146;
 Matches 645; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 QY 1 TCCGTTATCGCTAACACAGATGACCTACAAAGTTTATATGTCCAGGCACGGTCAATGGACAC 60
 DB 4 TCCGTTATCGCTAACACAGATGACCTACAAAGTTTATATGTCCAGGCACGGTCAATGGACAC 63
 QY 61 TACTTTGAGTTCGAGGCGATGGAAGAGAAAGCCTTACGAGGGGGAGCAGACGGTAAGG 120
 DB 64 TACTTTGAGTTCGAGGCGATGGAAGAGAAAGCCTTACGAGGGGGAGCAGACGGTAAGG 123

QY 121 CTGGCTGTCAACAAAGGGGGACCTCTGCTCCATTGCTTGGGATATTTTATCACCACAGTGT 180
 DB 124 CTCACCTGTCAACAAAGGGTGGACCTCTGCTCCATTGCTTGGGATATTTTATCACCACAGTCA 183
 QY 181 CAGTACGGAAGCATACCATTCACCAAGTACCTCAAGACATCCCTGACTATCTTAAAGCAG 240
 DB 184 CAGTACGGAAGCATACCATTCACCAAGTACCTCAAGACATCCCGGACTATGTAAAGCAG 243
 QY 241 TCATTCCCGGGGAGATATACATGGGAGAGGATCATGAATTTTGAAGATGGTGCAAGTGTGT 300
 DB 244 TCATTCCCGGAGGATATACATGGGAGAGGATCATGAATTTTGAAGATGGTGCAAGTGTGT 303
 QY 301 ACTGTACCAATGATTCAGCATCAAGGCAACTGTTTTCATCTACCATGTCAAGTTCTCT 360
 DB 304 ACTGTACCAATGATTCAGCATCAAGGCAACTGTTTTCATCTACCATGTCAAGTTCTCT 363
 QY 361 GGTTTGAACCTTTCTCCCAATGGACCTGTTATGACAGAAAGACACAGGGCTGGGAACCC 420
 DB 364 GGTTTGAACCTTTCTCCCAATGGACCTGTTATGACAGAAAGACACAGGGCTGGGAACCC 423
 QY 421 AACACTGAGCGTCTCTTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG 480
 DB 424 AACACTGAGCGTCTCTTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTG 483
 QY 481 AAGTTAGAAGGAGGTGCTCACTATTGTTGTAATTCAAATCTACTTACAAAGGCAAGGAAG 540
 DB 484 AAGTTAGAAGGAGGTGCTCACTATTGTTGTAATTCAAATCTACTTACAAAGGCAAGGAAG 543
 QY 541 CCTGTGAAGATGCCAGGGTATCACTATGTTTCAACGCAAACTGGATGTAAACCAATCACAAC 600
 DB 544 CCTGTGAAGATGCCAGGGTATCACTATGTTTCAACGCAAACTGGATGTAAACCAATCACAAC 603
 QY 601 AAGGATTACACTTCGGTTGAGCAGCGTGAATTTCCATTTCACGCAAACTTTTGGTGGCC 660
 DB 604 AAGGATTACACTTCGGTTGAGCAGCGTGAATTTCCATTTCACGCAAACTTTTGGTGGCC 663

Search completed: August 13, 2004, 17:09:51
 Job time : 410.5 secs

This Page Blank (uspto)

124 GCTGTACCAAGGCGGACCTCTGCCATTTGCTTGGATATTTTATCACCACAGTGTCAG 183
139 AAGGTAAACCAAGGGGGACCTTTGGCCATTTGCTTGGATATTTTGTGCCACACATTTTCAG 198
184 TAGGGAAGCATATACATTCACCAAGTACCTCTGAAGACATCCCTGACTATGTAAAGCAGTCA 243
199 TATGGAAGCAAGGTATATGTCAAGCATCTCTGCCACATACAGACTATAAAAGCTGTCA 258
244 TTCCCGGGAGATATACATGGGAGGAGATCATGAACCTTTGAAGATGTGTGAGTGTACT 303
259 TTTCTGAAGGATTTAAATGGGAAGGGTCATGAACCTTTGAAGACGGTGGCGTCTGTACT 318
304 GTGAGCAATGATTCACCATCCAGGCAACTGTTTTCATCTACCATGTCAAGTTCCTCTGGT 363
319 GTAAACCCAGGATTCAGTTTTCAGGATGGCTGTTCATCTACAAGGTCAAGTTCATTGGC 378
364 TTGAACCTTTCTCCCAATGCACTGTTATGAGAGAGACACAGGGCTGGGAACCCAAAC 423
379 GTGAACCTTTCTCTGATGGACCTGTTATGCAAAAGAGACAATGGGCTGGGAAGCCAGC 438
424 ACTGAGGCTCTTTTGACAGAGATGGAATGCTGTATAGGAACACTTTATGGCTCTGAAG 483
439 ACTGAGGCTTTGTATCTCTGATGGCGTGTGAAAGGAGAGATTCATAAGGCTCTGAAG 498
484 TTAGAAGGAGGTGTCTACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAGGAAGCCT 543
499 TTGAAGACGGTGTCTACTTACCTAGTTGAATTCAAACTATTTACATGGCAGGAAGCCT 558
544 GTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAAACAATCAACAAG 603
559 GTGACGTACCAAGGTACTACTATGTTGACTCCAACTGGATATAACAAGCCACAACAAA 618
604 GATTACACTTCGTTGACAGCGTGAATTTCCATTGACCAACACCTTTGGTC 657
619 GACTATACATCTGTTGACAGTATGAAGAAGCAAGGAGGACGCCACCATCTCTGTC 672

RESULT 2

US-10-314-936-3
; Sequence 3, Application US/10314936
; Publication No. US20040110225A1
; GENERAL INFORMATION:
; APPLICANT: Gibbs, Patrick D.L.
; APPLICANT: Carter, Robert W.
; APPLICANT: Schmale, Michael C.
; TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES
; FILE REFERENCE: 638,004
; CURRENT APPLICATION NUMBER: US/10/314,936
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: mutant red fluorescent protein
; NAME/KEY: CDS
; LOCATION: (1)..(711)
; OTHER INFORMATION:
US-10-314-936-3

Query Match 45.8%; Score 385.2; DB 17; Length 711;
Best Local Similarity 74.3%; Pred. No. 5.4e-95;
Matches 486; Conservative 0; Mismatches 168; Indels 0; Gaps 0;
4 GTTATCGCTAAACAGATGACCTACAAAGTTTATGTACGGCAGCGTCAATGGACACTAC 63
19 GTATCAAGAGTTTCATGAGTTTAAAGTTTCTGATGGAAGGAACGGTCAATGGCAGCAG 78
64 TTTGAGTCAAGGCGATGGAAGAAGCAAGCCCTTACGAGGGGAGCAGACGGTAAGGCTG 123

79 TTTGAAATAGAGCGCAAGGAGGAGGCCCATACGAAGGCCCAATACCGTAAAGCTT 138
124 GCTGTACCAAGGCGGACCTCTGCCATTTGCTTGGATATTTTATCACCACAGTGTCAG 183
139 AAGGTAAACCAAGGGGGACCTTTGGCCATTTGCTTGGATATTTTGTGCCACACATTTTCAG 198
184 TAGGGAAGCATATACATTCACCAAGTACCTCTGAAGACATCCCTGACTATGTAAAGCAGTCA 243
199 TATGGAAGCAAGGTATATGTCAAGCATCTCTGCCACATACAGACTATAAAAGCTGTCA 258
244 TTCCCGGGAGATATACATGGGAGGAGATCATGAACCTTTGAAGATGTGTGAGTGTACT 303
259 TTTCTGAAGGATTTAAATGGGAAGGGTCATGAACCTTTGAAGACGGTGGCGTCTGTACT 318
304 GTGAGCAATGATTCACCATCCAGGCAACTGTTTTCATCTACCATGTCAAGTTCCTCTGGT 363
319 GTAAACCCAGGATTCAGTTTTCAGGATGGCTGTTCATCTACAAGGTCAAGTTCATTGGC 378
364 TTGAACCTTTCTCCCAATGCACTGTTATGAGAGAGACACAGGGCTGGGAACCCAAAC 423
379 GTGAACCTTTCTCTGATGGACCTGTTATGCAAAAGAGACAATGGGCTGGGAAGCCAGC 438
424 ACTGAGGCTCTTTTGACAGAGATGGAATGCTGTATAGGAACACTTTATGGCTCTGAAG 483
439 ACTGAGGCTTTGTATCTCTGATGGCGTGTGAAAGGAGAGATTCATAAGGCTCTGAAG 498
484 TTAGAAGGAGGTGTCTACTATTTGTGTGAATTCAAATCTACTTACAAGGCAAGGAAGCCT 543
499 TTGAAGACGGTGTCTACTTACCTAGTTGAATTCAAACTATTTACATGGCAGGAAGCCT 558
544 GTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAAACAATCAACAAG 603
559 GTGACGTACCAAGGTACTACTATGTTGACTCCAACTGGATATAACAAGCCACAACAAA 618
604 GATTACACTTCGTTGACAGCGTGAATTTCCATTGACCAACACCTTTGGTC 657
619 GACTATACATCTGTTGACAGTATGAAGAAGCAAGGAGGACGCCACCATCTCTGTC 672

RESULT 3

US-10-006-922-9
; Sequence 9, Application US/10006922
; Publication No. US20020197676A1
; GENERAL INFORMATION:
; APPLICANT: Lukyanov, Sergey A.
; APPLICANT: Pradkov, Arcady F.
; APPLICANT: Labas, Yulii A.
; APPLICANT: Matz, Mikhail V.
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and
; TITLE OF INVENTION: Methods for Using the Same
; FILE REFERENCE: CLON-035CIP
; CURRENT APPLICATION NUMBER: US/10/006,922
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/120,330
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/457,898
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,144
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,477
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/457,556
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/444,338
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 850
; TYPE: DNA
; ORGANISM: Disosoma striata

US-10-006-922-9

Query Match 45.3%; Score 381.2; DB 14; Length 850;
Best Local Similarity 72.4%; Pred. No. 7.4e-94;
Matches 494; Conservative 0; Mismatches 188; Indels 0; Gaps 0;

QY 4 GTTATCGCTAAACAGATGACCTTACAAAGTTTATATGTCTCAGGCACCGTCAATGGACACTAC 63
DB 60 GTGATCAAGGAAGAAATGTTGATCGATCTTCTATCTGGAAGGAACGTTCAATGGGCACTAC 119

QY 64 TTTGAGGTGGAAGCGATGGAAGAAAGCCTTACGAGGGGAGACAGCGTAAGGCTG 123
DB 120 TTTGAAATAAAGGCAAGGAAGAAAGGAGCGCTATGAAAGGACCAATACCGTCAAGCTC 179

QY 124 GCTGTCAACAGGGGCGACCTCTGCCATTTGCTTGGGATATTTATCACCAAGTGTTCAG 183
DB 180 GAGGTTTCAAGGGTGGACCTCTGCCATTTGCTTGGGATATTTGTCGCCCAATTTTCAG 239

QY 184 TACGGAAGCAVACCAATTCAACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCA 243
DB 240 TATGGAACAAGGCATTTGTCCACCACCTGACGACATACCTGATTATCTAAAGCTGTCA 299

QY 244 TTCCCGGGGAGATATACATGGGAGGAGATCATGAACCTTTGAAGATGGTGCAGTGTACT 303
DB 300 TTTCCGGAGGATATACATGGGAGGAGCGTCCATGTCACCTTTGAAGACGGTGTGTGTGT 359

QY 304 GTCAGCAATGATCCAGCATCCAGGCAACTGTTTCACTACCATGTCTCAAGTCTCTGGT 363
DB 360 ATCCCAATGAATCAGTTTGACAGGCAACTGTTTCACTACCATGTCTCAAGTCTCTGGT 419

QY 364 TTGAACCTTCTCCCAATGGACCTGTTATGAGAGAGACACAGGGCTGGGAACCCCAAC 423
DB 420 TTGAACCTTCTCCCAATGGACCTGTTATGAGAGAGACACACTGGCTGGGACCGAGC 479

QY 424 ACTGAGCTCTCTTTCAGCAGATGGAATGCTGATAGGAACAACCTTTATGGCTCTGAAG 483
DB 480 ACTGAGCTTGTATCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 539

QY 484 TTGAAGAGGTGTCATATTTGTGTAATTCAACTACTTACAGGCAAGGAGCT 543
DB 540 GTGAAGAGGTGTCATATGATGACATTAATTAATTAATTAATTAATTAATTAATTAAT 599

QY 544 GTGAAGATGTCAGGGTATCACTATGTCAGCGCAAACTGGATGTAACCAATCAACAACAG 603
DB 600 GTAAAGATGTCAGGGTATCACTATGTCAGCACCAAACTGGTATTAAGAGCAACGACAA 659

QY 604 GATTACACTTCCTGAGCAGCGTGAATTTCCATTTGACGCAAACTTTGGTGCCTGC 663
DB 660 GAATTCATGAAGTTGAGGAGCATGAATCGCGTTGACGCGCCACCATCCGCTCCAAAGC 719

QY 664 TGTCTTTTCAGAGTCAAAATCAA 685
DB 720 CAATGAAGCTTAAGTAAGCAA 741

RESULT 4

US-10-332-733-22
; Sequence 22, Application US/10332733
; Publication No. US2004010656A1
; GENERAL INFORMATION:
; APPLICANT: Margaret Odenthal and Diana Jung
; TITLE OF INVENTION: Gene Expression, Genome Alteration And Reporter Expression
; FILE REFERENCE: 1472/68806
; CURRENT APPLICATION NUMBER: US/10/332,733
; CURRENT FILING DATE: 2003-06-10
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 22
; LENGTH: 666
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:

; OTHER INFORMATION: Description of artificial sequence: Red
; OTHER INFORMATION: Fluorescent Protein
US-10-332-733-22

Query Match 45.2%; Score 380.4; DB 17; Length 666;
Best Local Similarity 73.9%; Pred. No. 1.1e-93;
Matches 483; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 4 GTTATCGCTAAACAGATGACCTTACAAAGTTTATATGTCTCAGGCACCGTCAATGGACACTAC 63
DB 7 GTTATCAAGGAGTTTCATGAGGTTTAAAGGTTTCGATGGAAGGAACGTTCAATGGGCAAG 66

QY 64 TTTGAGGTGGAAGCGATGGAAGAAAGCCTTACGAGGGGAGACAGCGTAAGGCTG 123
DB 67 TTTGAAATAGAAAGCGGAGAGAGGAGGAGCCATACGAAGGCCCAATACCGTAAAGTT 126

QY 124 GCTGTCAACAGGGGCGACCTCTGCCATTTGCTTGGGATATTTATCACCAAGTGTTCAG 183
DB 127 AAGGTAAACCAAGGGGAGCCTTTCCATTTGCTTGGGATATTTTCTCACCAATTTTCAG 186

QY 184 TACGGAAGCATACCATTCACCAAGTACCCTGAAGACATCCCTGACTATGTAAAGCAGTCA 243
DB 187 TATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCAAGCTATAAAGAGCTGCA 246

QY 244 TTCCCGGGGAGATATACATGGGAGGAGTCAATGAACCTTTGAAGATGGTGCAGTGTACT 303
DB 247 TTTCTGAGGATTTAAATGGGAAAGGGTCAIGAACTTTTGAAGACGGTGGCTGCTTACT 306

QY 304 GTCAGCAATGATCCAGCATCCAGGCAACTGTTTCACTACCATGTCTCAAGTCTCTGGT 363
DB 307 GTAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCACTACAGGTCAGTTCAATTGGC 366

QY 364 TTGAACCTTCTCCCAATGGACCTGTTATGAGAGAGAGACACAGGGCTGGGAACCCCAAC 423
DB 367 GTGAACCTTCTCCCAATGGACCTGTTATGACAAAGAGACAAATGGCTGGGAGCCAGC 426

QY 424 ACTGAGCTCTCTTTCAGCAGATGGAATGCTGATAGGAACAACCTTTATGGCTCTGAAG 483
DB 427 ACTGAGCTTGTATCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 486

QY 484 TTGAAGAGGTGTCATATTTGTGTAATTCAACTACTTACAGGCAAGGAGCT 543
DB 487 CTGAAGAGCGTGTCTCATTTACTAGTTGAAATTCAAAGTATTTACATGGCAAGAGCT 546

QY 544 GTGAAGATGTCAGGGTATCACTATGTCAGCGCAAACTGGATGTAACCAATCAACAAG 603
DB 547 GTGCAAGTACCAAGGTTACTATGTTGACTCCAACTGGATATAACAAGCCCAACGAA 606

QY 604 GATTACACTTCCTGAGCAGCGTGAATTTCCATTTGACGCAAACTTTGGTGC 657
DB 607 GACTATACATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 660

RESULT 5

US-09-967-772-6
; Sequence 6, Application US/09967772
; Patent No. US2002016457A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSJEN, Roger
; APPLICANT: GONZALEZ, Jesus
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS
; FILE REFERENCE: REGEN1290-5
; CURRENT APPLICATION NUMBER: US/09/967,772
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 09/459,956
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: US 08/765,860
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: PCT/US96/09652
; PRIOR FILING DATE: 1996-06-06
; PRIOR APPLICATION NUMBER: US 08/481,977
; PRIOR FILING DATE: 1995-06-07

```
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp "red"
US-09-967-772-6

Query Match      45.2%; Score 380.4; DB 9; Length 678;
Best Local Similarity 73.9%; Pred. No. 1.1e-93;
Matches 483; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY  4 GTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGTCAGGACGGTCAATGGACACTAC 63
Db  19 GTTATCAAGGAGTTTCATGAGTTTAAGGTTTCGCGATGGAAGAACCGGTCAATGGCCAGAG 78
QY  64 TTTGAGTTCGAAGCGGATGAAAAGGAAAGCCTTACGAGGGGAGCAGACGGTAAGCGTG 123
Db  79 TTTGAATATAGAGCGGAGGAGGAGGAGGCCATACGAAGGCCCAATACCGTAAAGCTT 138
QY  124 GCTGTCTACCAAGGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGTGAG 183
Db  139 AAGGTAAACCAAGGGGAGCCTTTGCCATTTGCTTGGGATATTTTGTCCACCAATTTTCA 198
QY  184 TAACGAGAGCATACCAATTCACCAAGTACCCCTGAAGACATCCCTGACTATGTAAAGCAGTCA 243
Db  199 TATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACGAGACTATAAAAGCTGTCA 258
QY  244 TTCCCGGGGAGATATACATGGGAGAGGATCATGAACCTTTGAAGATGTGAGTGTGACT 303
Db  259 TTTCTCGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAAGACGGTGGCGTCTGTACT 318
QY  304 GTGAGCAATGATTCAGCATCCAGGCAACTGTTTTCATCTACCATGTCAAGTTCTCTGGT 363
Db  319 GTAACCCAGATTCAGTTTCCAGATGGCTGTTTCATCTACAAGGTCAGTTCAATTTGGC 378
QY  364 TTGAACCTTCTCCCAATGACCTGTTTATGAGAAAGAGACACAGGGCTGGAAACCCCAAC 423
Db  379 GTGAACCTTCTCCGATGAGCCCTGTTATGCAAAAGAGACAAATGGGCTGGGAGCCAGC 438
QY  424 ACTGAGCGTCTCTTTGACCGAGATGGAATGCTGATGAGAAACAACTTTATGGCTCTGAAG 483
Db  439 ACTGAGCGTTGTATCTCTCGTGTGCGGTGTTGAAAGAGAGATTCATAAGGCTCTGAAG 498
QY  484 TTGAAGAGCGTGTGCTACTATTTGTGTAATTCAAATCTACTTACAGGCGAGGAGGCT 543
Db  499 CTGAAGAGCGTGTGCTACTTACCTAGTTGAATTCAAAAGTATTTACATGGCAAGAGCCT 558
QY  544 GTGAAGATGCCAGGCTATCACTATGTTGACCGCAAACTGGATGTAAACCAATCAACAACAG 603
Db  559 GTGAGCTTACAGGCTACTACTATGTTGACTTCCAAACTGGATATAACAAGCCACACGAA 618
QY  604 GATTACACTTCCGTTGAGCAGCGTGAATTTCCATTTGCAACGCAACCTTTTGGTC 657
Db  619 GACTATACATCGTTGAGCAGTATGAAAGAACCGAGGAGCCACCATCTCTTTC 672
```

RESULT 6

```
US-10-132-067-3
; Sequence 3, Application US/10132067
; Publication No. US20030203355A1
; GENERAL INFORMATION:
; APPLICANT: Bradbury, Andrew
; APPLICANT: Zeytun, Ahmet
; APPLICANT: Waldo, Geoffrey
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Fluorobodies: Binding Ligands With Intrinsic
; FILE REFERENCE: 021362-000600US
; CURRENT APPLICATION NUMBER: US/10/132,067
; CURRENT FILING DATE: 2002-04-24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 3
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; OTHER INFORMATION: red fluorescent protein (dsRED)
; NAME/KEY: CDS
; LOCATION: (1)..(678)
; OTHER INFORMATION: dsRED
US-10-132-067-3
```

```
Query Match      45.2%; Score 380.4; DB 13; Length 678;
Best Local Similarity 73.9%; Pred. No. 1.1e-93;
Matches 483; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY  4 GTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGTCAGGACGGTCAATGGACACTAC 63
Db  19 GTTATCAAGGAGTTTCATGAGTTTAAGGTTTCGCGATGGAAGAACCGGTCAATGGCCAGAG 78
QY  64 TTTGAGTTCGAAGCGGATGAAAAGGAAAGCCTTACGAGGGGAGCAGACGGTAAGCGTG 123
Db  79 TTTGAATATAGAGCGGAGGAGGAGGAGGCCATACGAAGGCCCAATACCGTAAAGCTT 138
QY  124 GCTGTCTACCAAGGCGGACCTCTGCCATTTGCTTGGGATATTTTATCACCACAGTGTGAG 183
Db  139 AAGGTAAACCAAGGGGAGCCTTTGCCATTTGCTTGGGATATTTTGTCCACCAATTTTCA 198
QY  184 TAACGAGAGCATACCAATTCACCAAGTACCCCTGAAGACATCCCTGACTATGTAAAGCAGTCA 243
Db  199 TATGGAAGCAAGGTATATGTCAAGCACCCCTGCCGACATACGAGACTATAAAAGCTGTCA 258
QY  244 TTCCCGGGGAGATATACATGGGAGAGGATCATGAACCTTTGAAGATGTGAGTGTGACT 303
Db  259 TTTCTCGAAGGATTTAAATGGGAAAGGTCATGAACCTTTGAAGACGGTGGCGTCTGTACT 318
QY  304 GTGAGCAATGATTCAGCATCCAGGCAACTGTTTTCATCTACCATGTCAAGTTCTCTGGT 363
Db  319 GTAACCCAGATTCAGTTTCCAGATGGCTGTTTTCATCTACAAGGTCAGTTCAATTTGGC 378
QY  364 TTGAACCTTCTCCCAATGACCTGTTTATGAGAAAGAGACACAGGGCTGGAAACCCCAAC 423
Db  379 GTGAACCTTCTCCGATGAGCCCTGTTATGCAAAAGAGACAAATGGGCTGGGAGCCAGC 438
QY  424 ACTGAGCGTCTCTTTGACCGAGATGGAATGCTGATGAGAAACAACTTTATGGCTCTGAAG 483
Db  439 ACTGAGCGTTGTATCTCTCGTGTGCGGTGTTGAAAGAGAGATTCATAAGGCTCTGAAG 498
QY  484 TTGAAGAGCGTGTGCTACTATTTGTGTAATTCAAATCTACTTACAGGCGAGGAGGCT 543
Db  499 CTGAAGAGCGTGTGCTACTTACCTAGTTGAATTCAAAAGTATTTACATGGCAAGAGCCT 558
QY  544 GTGAAGATGCCAGGCTATCACTATGTTGACCGCAAACTGGATGTAAACCAATCAACAACAG 603
Db  559 GTGAGCTTACAGGCTACTACTATGTTGACTTCCAAACTGGATATAACAAGCCACACGAA 618
QY  604 GATTACACTTCCGTTGAGCAGCGTGAATTTCCATTTGCAACGCAACCTTTTGGTC 657
Db  619 GACTATACATCGTTGAGCAGTATGAAAGAACCGAGGAGCCACCATCTCTTTC 672
```

RESULT 7

```
US-10-006-922-11
; Sequence 11, Application US/10006922
; Publication No. US20020197676A1
; GENERAL INFORMATION:
; APPLICANT: Lukanov, Sergey A.
; APPLICANT: Fradkov, Arcady F.
; APPLICANT: Labas, Yulii A.
; APPLICANT: Matz, Mikhail V.
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and
; METHODS FOR USING THE SAME
```

```
; FILE REFERENCE: CLON-035CIP
; CURRENT APPLICATION NUMBER: US/10/006,922
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/120,330
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/457,898
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,144
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/458,477
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/457,556
; PRIOR FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: 09/444,338
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma species
US-10-006-922-11

Query Match      45.2%; Score 380.4; DB 14; Length 678;
Best Local Similarity 73.9%; Pred. No. 1.1e-93;
Matches 483; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY      4 GTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGTCAGCAGCGTCAATGGACACTAC 63
Db      19 GTTATCAAGAGTTTCATGAGGTTTAAAGTTTGCATGGAAGAACGGTCAATGGGCACGAG 78

QY      64 TTTGAGGTCCGAAGCGCATGAAAGAGGAAGCCTTACGAGGGGGAGCAGAGCGTAAAGCTG 123
Db      79 TTTGAATATAGAGCGAGAGAGGGGAGGCATACGAGGCCACATACCGTAAAGCTT 138

QY      124 GCTGTACCAAGGCGGACCTTGCCATTTGCTTGGGATATTTTATCACCACAGTGTGAG 183
Db      139 AAGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACCAATTTTCAG 198

QY      184 TAGGAGCATACCATTCACCAAGTACCCCTGAAGACATCCCTGACTATGTAAAGCAGTCA 243
Db      199 TATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCAAGCTATAAAAGCTGCA 258

QY      244 TTCCCGGGGAGATATACATGGGAGGATCATGAACCTTTGAAGATGCTGCAAGTGTGACT 303
Db      259 TTTCTCGAAGGATTTAAATGGGAAGGGTCAATGAACCTTTGAAGACGGTGGCGTCTGACT 318

QY      304 GTGAGCAATGATTCAGCATCCAGGCAACTGTTTCACTACCATGTCATGTCATGTCCTGGT 363
Db      319 GTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAGTTCATTGGC 378

QY      364 TTGAACCTTCTCCCAATGGACCTGTATGACAGAGAGACACAGGGCTGGGAACCCAAC 423
Db      379 GTGAACCTTCTCCGATGGACCTGTATGCAAAAGAGACAATGGGCTGGGAAGCCAGC 438

QY      424 ACTGAGCGTCTCTTTGACAGAGATGGAATGCTGATAGAAACAACTTTATGGCTCTGAAG 483
Db      439 ACTGAGCGTTTGTATCTCTGATGGCGTGTGAAGAGAGATTCATAGGCTCTGAAG 498

QY      484 TTGAAGAGGAGTGTCTACTATTTGTGTAATTCAAATCTACTTCAAGGCAAGGAAGCCT 543
Db      499 CTGAAGAGCGGTGTCATTACCTAGTTGAAATTCAAAAGTATTTACATGGCAAGGAAGCCT 558

QY      544 GTGAGATGCCAGGTATCACTATGTTGACCGGAAACTGATGTAAACCAATCAACAAG 603
Db      559 GTGAGATGCCAGGTATCACTATGTTGACTCCAAACTGGATATTAACAAGCCACCAAGAA 618

QY      604 GATTACACTTCCGTTGAGCAGCGTGAATTTCCATTTGCAAGCAACCTTTGGTC 657
Db      619 GACTATACATCGTTGAGCAGTATGAAGAACCGGAGGACGCCACCATCTGTTTC 672
```

```
US-10-121-258-2
; Sequence 2, Application US/10121258
; Publication No. US20030059835A1
; GENERAL INFORMATION:
; APPLICANT: Tsen, Robert
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; FILE REFERENCE: UC083.1CP2CP1
; CURRENT APPLICATION NUMBER: US/10/121,258
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(678)
; OTHER INFORMATION: wild-type DsRed
US-10-121-258-2

Query Match      45.2%; Score 380.4; DB 15; Length 678;
Best Local Similarity 73.9%; Pred. No. 1.1e-93;
Matches 483; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY      4 GTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGTCAGCAGCGTCAATGGACACTAC 63
Db      19 GTTATCAAGAGTTTCATGAGGTTTAAAGTTTGCATGGAAGAACGGTCAATGGGCACGAG 78

QY      64 TTTGAGGTCCGAAGCGCATGAAAGAGGAAGCCTTACGAGGGGGAGCAGAGCGTAAAGCTG 123
Db      79 TTTGAATATAGAGCGAGAGAGGGGAGGCATACGAGGCCACATACCGTAAAGCTT 138

QY      124 GCTGTACCAAGGCGGACCTTGCCATTTGCTTGGGATATTTTATCACCACAGTGTGAG 183
Db      139 AAGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTCAACCAATTTTCAG 198

QY      184 TAGGAGCATACCATTCACCAAGTACCCCTGAAGACATCCCTGACTATGTAAAGCAGTCA 243
Db      199 TATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCAAGCTATAAAAGCTGCA 258

QY      244 TTCCCGGGGAGATATACATGGGAGGATCATGAACCTTTGAAGATGCTGCAAGTGTGACT 303
Db      259 TTTCTCGAAGGATTTAAATGGGAAGGGTCAATGAACCTTTGAAGACGGTGGCGTCTGACT 318

QY      304 GTGAGCAATGATTCAGCATCCAGGCAACTGTTTCACTACCATGTCATGTCATGTCCTGGT 363
Db      319 GTAAACCCAGGATTCAGTTTGCAGGATGGCTGTTTCATCTACAAGGTCAAGTTCATTGGC 378

QY      364 TTGAACCTTCTCCCAATGGACCTGTATGACAGAGAGACACAGGGCTGGGAACCCAAC 423
Db      379 GTGAACCTTCTCCGATGGACCTGTATGCAAAAGAGACAATGGGCTGGGAAGCCAGC 438

QY      424 ACTGAGCGTCTCTTTGACAGAGATGGAATGCTGATAGAAACAACTTTATGGCTCTGAAG 483
Db      439 ACTGAGCGTTTGTATCTCTGATGGCGTGTGAAGAGAGATTCATAGGCTCTGAAG 498

QY      484 TTGAAGAGGAGTGTCTACTATTTGTGTAATTCAAATCTACTTCAAGGCAAGGAAGCCT 543
Db      499 CTGAAGAGCGGTGTCATTACCTAGTTGAAATTCAAAAGTATTTACATGGCAAGGAAGCCT 558

QY      544 GTGAGATGCCAGGTATCACTATGTTGACCGGAAACTGATGTAAACCAATCAACAAG 603
Db      559 GTGAGATGCCAGGTATCACTATGTTGACTCCAAACTGGATATTAACAAGCCACCAAGAA 618

QY      604 GATTACACTTCCGTTGAGCAGCGTGAATTTCCATTTGCAAGCAACCTTTGGTC 657
```

Db 619 GACTATACAATCGTTGAGCAGTATGAAGAAGACCGAGGAGCGCCACCACCTCTGTTTC 672

RESULT 9

US-10-335-517-6

; Sequence 6, Application US/10335517

; Publication No. US20030207248A1

; GENERAL INFORMATION:

; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA

; APPLICANT: TSJEN, Roger

; APPLICANT: GONZALEZ, Jesus

; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS

; FILE REFERENCE: REGEN1290-5

; CURRENT APPLICATION NUMBER: US/10/335,517

; CURRENT FILING DATE: 2002-12-31

; PRIOR APPLICATION NUMBER: US/09/967,772

; PRIOR FILING DATE: 2001-09-28

; PRIOR APPLICATION NUMBER: US 09/459,956

; PRIOR FILING DATE: 1999-12-13

; PRIOR APPLICATION NUMBER: US 08/765,860

; PRIOR FILING DATE: 1996-12-19

; PRIOR APPLICATION NUMBER: PCT/ US96/09652

; PRIOR FILING DATE: 1996-06-06

; PRIOR APPLICATION NUMBER: US 08/481,977

; PRIOR FILING DATE: 1995-06-07

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 6

; TYPE: DNA

; LENGTH: 678

; ORGANISM: Discosoma sp "red"

US-10-335-517-6

Query Match 45.2%; Score 380.4; DB 16; Length 678;

Best Local Similarity 73.9%; Pred. No. 1.1e-93;

Matches 483; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 4 GTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGTCAGGCGAGCGTCAATGGACACTAC 63

Db 19 GTTATCAAGGAGTTTCATGAGTTTAAAGTTTCGATGAAGGAAAGCGTCAATGGGACGAG 78

QY 64 TTTGAGTTCGAAGCGATGAAAGAAAGAAAGCTTACGAGGGGAGGAGAGCGTTAAGCTG 123

Db 79 TTTGAATAGAAAGCGAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 138

QY 124 GCTGTCAACCAAGGCGGACCTCTGCCATTTGCTTTGGATATTTTATCACCACAGTGTGAG 183

Db 139 AAGGTAAACCAAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 198

QY 184 TACGGAAGCATACATTCACCAAGTACCCCTGAGACATCCCTGACTATGTAAAGCAGTCA 243

Db 199 TATGGAAGCAAGGTATATGTCAAGCACCCTGCCACATACAGGAGGAGGAGGAGGAGGAGGAG 258

QY 244 TTCCCGGGAGATATACATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 303

Db 319 GTAAACCAAGATTCAGTTTCAGGATGGCTGTTTCATCTACAGGTCAAGTTCAATTTGGC 378

QY 364 TTGAACCTTCTCCCAATGACCTGTTATGAGAAAGAGACACAGGCGTGGGAGCCCAAC 423

Db 379 GTGAACCTTCTCCCAATGACCTGTTATGAGAAAGAGACAAATGGGCTGGGAGCCAGC 438

QY 424 ACTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGGAACAACTTTATGGCTCTGAAG 483

Db 439 ACTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGGAACAACTTTATGGCTCTGAAG 498

QY 484 TTGAAGAGGAGTGTGCTCACTATTTGTTGTAATCAAACTCTACTTACAAGGCAAGGAGCCT 543

Db 499 CTGAAGAGCGTGTGCTCACTATTTGTTGTAATCAAACTCTACTTACAAGGCAAGGAGCCT 558

QY 544 GTGAGATGCCAGGATACATATGTTGACCGCAAACTGGATGTAACCAATCAACAAG 603

Db 559 GTGAGCTACCAAGGTTACTATGTTGACTCCAAACTGGATATAACAAGCCACAAGCAA 618

QY 604 GATTACACTTCGTTGAGCAGCGTGAAATTTCCATTCGACGCAAACTTTGGTC 657

Db 619 GACTATACAATCGTTGAGCAGTATGAAGAAGACCGAGGAGCGCCACCACCTCTGTTTC 672

RESULT 10

US-10-334-288-6

; Sequence 6, Application US/10334288

; Publication No. US20040002123A1

; GENERAL INFORMATION:

; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA

; APPLICANT: TSJEN, Roger

; APPLICANT: GONZALEZ, Jesus

; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY OPTICAL METHODS

; FILE REFERENCE: REGEN1290-5

; CURRENT APPLICATION NUMBER: US/10/334,288

; CURRENT FILING DATE: 2002-12-31

; PRIOR APPLICATION NUMBER: US/09/967,772

; PRIOR FILING DATE: 2001-09-28

; PRIOR APPLICATION NUMBER: US 09/459,956

; PRIOR FILING DATE: 1999-12-13

; PRIOR APPLICATION NUMBER: US 08/765,860

; PRIOR FILING DATE: 1996-12-19

; PRIOR APPLICATION NUMBER: PCT/ US96/09652

; PRIOR FILING DATE: 1996-06-06

; PRIOR APPLICATION NUMBER: US 08/481,977

; PRIOR FILING DATE: 1995-06-07

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 6

; TYPE: DNA

; LENGTH: 678

; ORGANISM: Discosoma sp "red"

US-10-334-288-6

Query Match 45.2%; Score 380.4; DB 16; Length 678;

Best Local Similarity 73.9%; Pred. No. 1.1e-93;

Matches 483; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 4 GTTATCGCTAAACAGATGACCTACAAAGTTTATATGTGTCAGGCGAGCGTCAATGGACACTAC 63

Db 19 GTTATCAAGGAGTTTCATGAGTTTAAAGTTTCGATGAAGGAAAGCGTCAATGGGACGAG 78

QY 64 TTTGAGTTCGAAGCGATGAAAGAAAGAAAGCTTACGAGGGGAGGAGAGCGTTAAGCTG 123

Db 79 TTTGAATAGAAAGCGAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 138

QY 124 GCTGTCAACCAAGGCGGACCTCTGCCATTTGCTTTGGATATTTTATCACCACAGTGTGAG 183

Db 139 AAGGTAAACCAAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 198

QY 184 TACGGAAGCATACATTCACCAAGTACCCCTGAGACATCCCTGACTATGTAAAGCAGTCA 243

Db 199 TATGGAAGCAAGGTATATGTCAAGCACCCTGCCACATACAGGAGGAGGAGGAGGAGGAGGAG 258

QY 244 TTCCCGGGAGATATACATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 303

Db 259 TTTCCCTGAAGATTTAAATGGGAAAGGTCATGAACCTTTGAAAGCGTGGGCTGCTTACT 318

QY 304 GTGAGCAATGATTCAGATCCAGGCAACTGTTTCATCTACCATGTCAGATTCCTCTGGT 363

Db 319 GTAAACCAAGATTCAGTTTCAGGATGGCTGTTTCATCTACAGGTCAAGTTCAATTTGGC 378

QY 364 TTGAACCTTCTCCCAATGACCTGTTATGAGAAAGAGACACAGGCGTGGGAGCCCAAC 423

Db 379 GTGAACCTTCTCCCAATGACCTGTTATGAGAAAGAGACAAATGGGCTGGGAGCCAGC 438

QY 424 ACTGAGCGTCTCTTTGACGAGATGGAATGCTGATAGGAACAACTTTATGGCTCTGAAG 483

Db 439 ACTGAGCGTTTGTATCTCTCGTGTATGGCGTGTGAAAGGAGAGATTTCATAAGGCTCTGAAG 498
Qy 484 TTAGAGAGAGTGTCTACATATTTGTGTGAATTCAAATCTACTTACAGGCGAAGAGCGCT 543
Db 499 CTGAAAGACGGTGGTTCATTACTAGTTGAATTCAAAGTATTTACATGGCAAGAGCGCT 558
Qy 544 GTGAGATGCGCAGGGTATCATTCTTTCACCGCAAACTGGATGTAAACCAATCACAAAG 603
Db 559 GTGAGCTACAGGGTACTACTATGTTGACCTCCAACTGGATGTAAAGCCACAAAGAA 618
Qy 604 GATTACACTCCGTTGACAGCGTGAATTTCCATTTGCACGCAAACTTTGGTC 657
Db 619 GACTATACATCGTTGACAGTATGAAGAACCGAGGAGCGCCACCATCTGTTTC 672

RESULT 11

US-09-999-745-66
; Sequence 66, Application US/09999745
; Patent No. US20020157120A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Baird, Geoffrey
; FILE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS
; TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS
; FILE REFERENCE: REGEN1470-1
; CURRENT APPLICATION NUMBER: US/09/999,745
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 09/316,920
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 66
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
US-09-999-745-66

Query Match 45.2%; Score 380.4; DB 9; Length 859;
Best Local Similarity 73.9%; Pred. No. 1.2e-93;
Matches 483; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

Qy 4 GTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGACGGTCAATGGACACTAC 63
Db 72 GTTATCAAGGAGTTTCATGAGGTTTAAAGTTCCGATGGAAGAACCGTCAATGGGACGAG 131
Qy 64 TTTGAGGTTCGAAGGCGATGGAAGGAAAGGCTTACGAGGGGGAGCAGACCGTAAAGGCTG 123
Db 132 TTTGAAATAGAAAGCGAAGGAGAGGGGAGGCCATACGAAGGCCCAATACCGTAAAGCTT 191
Qy 124 GCTGTACCAAGGGCGACCTCTGCCATTTGCTTGGGATATTTTATCACCAAGTGTGAG 183
Db 192 AAGGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTACCAACAATTTTCCAG 251
Qy 184 TACGGAAGCATACCATTCACCAAGTACCCCTGAAGACATCCCTGACTATGTAAGCAGTCA 243
Db 252 TATGGAAGCAAGGTATATGTCAAGCACCTCCGACATACCAGACTATAAAGAGCTGTCA 311
Qy 244 TTCCCGGGGAGATATACATGGGAGGAGTATGAATTTTGAAGATGGTGCAGTGTGACT 303
Db 312 TTTCTCTGAAGGATTTAAATGGAAGGAGGTATGAACTTTGAGACGGTGGCGTCTGTACT 371
Qy 304 GTGAGCAATGATTCAGGATCCAGGCAACTGTTTTCATCTACATGTCACCAAGTCTCTGGT 363
Db 372 GTAAACCCAGGATTCAGTTCAGGATGGCTGTTTTCATCTCAAGGTCAAGTTCATTGGC 431
Qy 364 TTGAACCTTCTCCCAATGGACCTGTTATGAGAGAGAGACACAGGGCTGGGACCCAC 423
Db 432 GTGAACCTTCTCCGATGGACCTGTTATGCAAAAGAGACAATGGGGCTGGGAGCCAGC 491
Qy 424 ACTGAGCGTCTCTTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAG 483

Db 492 ACTGAGCGTTTGTATCTCTCGTGTATGGCGTGTGAAAGGAGAGATTTCATAAGGCTCTGAAG 551
Qy 484 TTAGAGAGAGTGTCTACATATTTGTGTGAATTCAAATCTACTTACAGGCGAAGAGCGCT 543
Db 552 CTGAAAGACGGTGGTTCATTACTAGTTGAATTCAAAGTATTTACATGGCAAGAGCGCT 611
Qy 544 GTGAGATGCGCAGGGTATCATTCTTTCACCGCAAACTGGATGTAAACCAATCACAAAG 603
Db 612 GTGAGCTACAGGGTACTACTATGTTGACCTCCAACTGGATGTAAAGCCACAAAGAA 671
Qy 604 GATTACACTCCGTTGACAGCGTGAATTTCCATTTGCACGCAAACTTTGGTC 657
Db 672 GACTATACATCGTTGACAGTATGAAGAACCGAGGAGCGCCACCATCTGTTTC 725

RESULT 12

US-09-866-538-11
; Sequence 11, Application US/09866538
; Publication No. US20030032088A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
; FILE REFERENCE: REGEN1530-2
; CURRENT APPLICATION NUMBER: US/09/866,538
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
US-09-866-538-11

Query Match 45.2%; Score 380.4; DB 10; Length 859;
Best Local Similarity 73.9%; Pred. No. 1.2e-93;
Matches 483; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

Qy 4 GTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAGGACGGTCAATGGACACTAC 63
Db 72 GTTATCAAGGAGTTTCATGAGGTTTAAAGTTCCGATGGAAGAACCGTCAATGGGACGAG 131
Qy 64 TTTGAGGTTCGAAGGCGATGGAAGGAAAGGCTTACGAGGGGGAGCAGACCGTAAAGGCTG 123
Db 132 TTTGAAATAGAAAGCGAAGGAGAGGGGAGGCCATACGAAGGCCCAATACCGTAAAGCTT 191
Qy 124 GCTGTACCAAGGGCGACCTCTGCCATTTGCTTGGGATATTTTATCACCAAGTGTGAG 183
Db 192 AAGGTAAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTACCAACAATTTTCCAG 251
Qy 184 TACGGAAGCATACCATTCACCAAGTACCCCTGAAGACATCCCTGACTATGTAAGCAGTCA 243
Db 252 TATGGAAGCAAGGTATATGTCAAGCACCTCCGACATACCAGACTATAAAGAGCTGTCA 311
Qy 244 TTCCCGGGGAGATATACATGGGAGGAGTATGAATTTTGAAGATGGTGCAGTGTGACT 303
Db 312 TTTCTCTGAAGGATTTAAATGGAAGGAGGTATGAACTTTGAGACGGTGGCGTCTGTACT 371
Qy 304 GTGAGCAATGATTCAGGATCCAGGCAACTGTTTTCATCTACATGTCACCAAGTCTCTGGT 363
Db 372 GTAAACCCAGGATTCAGTTCAGGATGGCTGTTTTCATCTCAAGGTCAAGTTCATTGGC 431
Qy 364 TTGAACCTTCTCCCAATGGACCTGTTATGAGAGAGAGACACAGGGCTGGGACCCAC 423
Db 432 GTGAACCTTCTCCGATGGACCTGTTATGCAAAAGAGACAATGGGGCTGGGAGCCAGC 491
Qy 424 ACTGAGCGTCTCTTTTGCACGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAG 483

Db 492 ACTGAGCGTTTGTATCTCTCGTATGCGGTGTGAAAGGAGAGATTCAATAGGCTCTGAAG 551
Qy 484 TTAGAGGAGGTGGTCACTATTGTGTGAATTCAAATCTACTTACAGGCAAGGAGCCT 543
Db 552 CTGAAAGACGGTGGTCACTACCTAGTTGAATTCAAAAGTATTACATGGCAAGAGCCT 611
Qy 544 GTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCAACAAG 603
Db 612 GTGAGCTACAGGGTACTACTATGTTGACTCCAACTGGATATAACAAGCCACAAGAA 671
Qy 604 GATTACATCTCCGTTGAGCAGCGTGAATTTCCATTGCGCAACACCTTTGGTC 657
Db 672 GACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGAGCGCCACCATCTGTTTC 725

RESULT 13

US-09-794-308-11
; Sequence 11, Application US/09794308
; Publication No. US20030170911A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSJEN, Roger
; APPLICANT: ZACHARIAS, David
; APPLICANT: BAIRD, Geoffrey
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
; FILE REFERENCE: REGEN1530
; CURRENT APPLICATION NUMBER: US/09/794,308
; CURRENT FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
US-09-794-308-11

Query Match 45.2%; Score 380.4; DB 10; Length 859;
Best Local Similarity 73.9%; Pred. No. 1.2e-93;
Matches 483; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
Qy 4 GTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAAGCAGCGTCAATGGACACTAC 63
Db 72 GTTATCAAGGAGTTTCATGAGGTTTAAAGTTTCGATGGAAGAACCGTCAATGGGACGAG 131
Qy 64 TTTGAGTTCGAAGCGGATGGAAGGAAGCCTTACGAGGGGAGCAGCGTAAAGCTG 123
Db 132 TTTGAAATAGAAAGCGAAGGAGAGGGGAGGCCATACGAAGGCCCAATACCGTAAAGCTT 191
Qy 124 GCTGTCAACCAAGGGGACCTCTGCCATTTCCTTGGGATATTTTATCACCACAGTGTGAG 183
Db 192 AAGTAAACCAAGGGGAGCCTTTGCCATTTCCTTGGGATATTTTGTCAACCAATTTGAG 251
Qy 184 TACGGAAGCATACCAATTCAACCAAGTACCGTGAAGACATCCCTGACTATGTAAGCAGTCA 243
Db 252 TATGGAAGCAAGGTATATGTCAAGCACCTCCGACATACCAAGTATAAAGAGCTGTCA 311
Qy 244 TTCCCGGGAGATATACATGGGAGGATCATGAATTTGAAGATGGTGCAGTGTGACT 303
Db 312 TTTCTTGAAGATTTAAATGGGAAGGGTCAAGACTTTTGAAGACGGTGGCGTCTGTTACT 371
Qy 304 GTCAAGCAATGATTCAGCATCAAGGCAACTGTGTTTCATCTACCTATGTCAAGTTCCTTGGT 363
Db 372 GTAACCCAGGATTCAGTTTCAGGATGGCTGTTTCATCTACAGGTCAAGTTCATTGGC 431
Qy 364 TTGAACCTTCTCCCAATGGACCTGTTATGAGAGAGAGACACAGGGCTGGAAACCAAC 423
Db 432 GTGAACCTTCTCCGATGGACCTGTTATGCAAAAGAGACAAATGGGCTGGGAAGCCAGC 491
Qy 424 ACTGAGCGTCTCTTTGACAGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAG 483

Db 492 ACTGAGCGTTTGTATCTCTCGTATGCGGTGTGAAAGGAGAGATTCAATAGGCTCTGAAG 551
Qy 484 TTAGAGGAGGTGGTCACTATTGTGTGAATTCAAATCTACTTACAGGCAAGGAGCCT 543
Db 552 CTGAAAGACGGTGGTCACTACCTAGTTGAATTCAAAAGTATTACATGGCAAGAGCCT 611
Qy 544 GTGAAGATGCCAGGGTATCACTATGTTGACCGCAAACTGGATGTAACCAATCAACAAG 603
Db 612 GTGAGCTACAGGGTACTACTATGTTGACTCCAACTGGATATAACAAGCCACAAGAA 671
Qy 604 GATTACATCTCCGTTGAGCAGCGTGAATTTCCATTGCGCAACACCTTTGGTC 657
Db 672 GACTATACAATCGTTGAGCAGTATGAAAGAACCGAGGAGCGCCACCATCTGTTTC 725

RESULT 14

US-09-865-291-11
; Sequence 11, Application US/09865291
; Publication No. US20030186229A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSJEN, Roger
; APPLICANT: TING, Alice
; APPLICANT: ZHANG, Jin
; TITLE OF INVENTION: EMISSION RATIONETRIC INDICATORS OF PHOSPHORYLATION
; FILE REFERENCE: REGEN1550
; CURRENT APPLICATION NUMBER: US/09/865,291
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
; LENGTH: 859
; TYPE: DNA
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
US-09-865-291-11

Query Match 45.2%; Score 380.4; DB 10; Length 859;
Best Local Similarity 73.9%; Pred. No. 1.2e-93;
Matches 483; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
Qy 4 GTTATCGCTAAACAGATGACCTACAAAGTTTATATGTCAAGCAGCGTCAATGGACACTAC 63
Db 72 GTTATCAAGGAGTTTCATGAGGTTTAAAGTTTCGATGGAAGAACCGTCAATGGGACGAG 131
Qy 64 TTTGAGTTCGAAGCGGATGGAAGGAAGCCTTACGAGGGGAGCAGCGTAAAGCTG 123
Db 132 TTTGAAATAGAAAGCGAAGGAGAGGGGAGGCCATACGAAGGCCCAATACCGTAAAGCTT 191
Qy 124 GCTGTCAACCAAGGGGACCTCTGCCATTTCCTTGGGATATTTTATCACCACAGTGTGAG 183
Db 192 AAGTAAACCAAGGGGAGCCTTTGCCATTTCCTTGGGATATTTTGTCAACCAATTTGAG 251
Qy 184 TACGGAAGCATACCAATTCAACCAAGTACCGTGAAGACATCCCTGACTATGTAAGCAGTCA 243
Db 252 TATGGAAGCAAGGTATATGTCAAGCACCTCCGACATACCAAGTATAAAGAGCTGTCA 311
Qy 244 TTCCCGGGAGATATACATGGGAGGATCATGAATTTGAAGATGGTGCAGTGTGACT 303
Db 312 TTTCTTGAAGATTTAAATGGGAAGGGTCAAGACTTTTGAAGACGGTGGCGTCTGTTACT 371
Qy 304 GTCAAGCAATGATTCAGCATCAAGGCAACTGTGTTTCATCTACCTATGTCAAGTTCCTTGGT 363
Db 372 GTAACCCAGGATTCAGTTTCAGGATGGCTGTTTCATCTACAGGTCAAGTTCATTGGC 431
Qy 364 TTGAACCTTCTCCCAATGGACCTGTTATGAGAGAGAGACACAGGGCTGGAAACCAAC 423
Db 432 GTGAACCTTCTCCGATGGACCTGTTATGCAAAAGAGACAAATGGGCTGGGAAGCCAGC 491
Qy 424 ACTGAGCGTCTCTTTGACAGAGATGGAATGCTGATAGGAAACAACTTTATGGCTCTGAAG 483

Db 492 ACTGAGCGTTTGTATCTCGTGATGGCGTGTGAAAGGAGAGATTCTAATGAGCTCTGAAG 551
Qy 484 TTAGAAGAGGTGTCTACTATTTGTGTGAATTCAAATCTACTTCAAGGCGAAGGAGCCT 543
Db 552 CTGAAGACGCGTGTCTACTTACCTAGTTGAATTCAAAAGTATTTACATGGCAAGAAGCCT 611
Qy 544 GTGAAGATGCCAGGTTATCACTATGTTGACCGCAAACTGGATGTAACCAATCAACAAG 603
Db 612 GTGACGTACCAAGGTTACTATGTTGACTCCAAACTGGATATACAAGCCACAAGAA 671
Qy 604 GATTACACTTCCGTTGAGCAGCGTGAATTTCCATTGACGCAAAACCTTTTGGTC 657
Db 672 GACTATACATCGTTGAGCAGTATGAAAGAACCGGAGGAGCCACCATTCTGTC 725

RESULT 15

US-10-433-640-12
; Sequence 12, Application US/10433640
; Publication No. US20040115792A1
; GENERAL INFORMATION:
; APPLICANT: Lichtenberg-Frate, Hella
; TITLE OF INVENTION: YEAST STRAIN FOR TESTING THE GENOTOXICITY AND CYTOTOXICITY OF
; TITLE OF INVENTION: COMPLEX ENVIRONMENTAL CONTAMINATIONS
; FILE REFERENCE: 1487/3
; CURRENT APPLICATION NUMBER: US/10/433,640
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/14610
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: DE 10061872.3
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 12
; LENGTH: 859
; TYPE: DNA
; ORGANISM: *Discosoma* sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (54)..(731)
US-10-433-640-12

Query Match 45.2%; Score 380.4; DB 17; Length 859;
Best Local Similarity 73.9%; Pred. No. 1.2e-93;
Matches 483; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
Qy 4 GTTATCGCTAAACAGATGACCTCAAGTTTATATGTCAGGCAGCGTCAATGGACACTAC 63
Db 72 GTTATCAAGGAGTTTCATGAGGTTTAAAGTTCCGATGGAAGAACGGTCAATGGGCACGAG 131
Qy 64 TTTGAGGTGCGAAGCGATGGAAGAAAGGAAAGCCTTACGAGGGGAGCGACCGTAAAGCGTG 123
Db 132 TTGGAATAGAAAGCGAAGAGAGGGAGGCCATACGAAGGCCCAATACCCGTAAAGCTT 191
Qy 124 GCTGTACCAAGGGGCGACCTCTGCCATTTGCTGGGATATTTTATCACCACAGTGTGAC 183
Db 192 AAGGTAACCAAGGGGGACCTTTGCCATTTGCTTGGGATATTTTGTACACCAATTTGAG 251
Qy 184 TACGGAAGCATACCATTCACCAAGTACCCCTGGAAGACATCCCTGACTATGTAAAGAGTCA 243
Db 252 TATGGAAGCAAGGTATATGTCAAGCACCCCTGCCACATACCAAGCCATATAAAAAGCTGCA 311
Qy 244 TTCCCGGGAGATATACATGGAGAGGATCATGAATTTGAAGTGGTGCAGTGTGACT 303
Db 312 TTTCTGAGGATTTAAATGGGAAGGGTCATGAACCTTTGAAGACGGTGGCGTCGTACT 371
Qy 304 GTCAGCAATGATCCAGCATCAAGGCAACTGTTTTCATCTACCATGTCAAGTTCCTGGT 363
Db 372 GTAACCCAGGATTCAGTTTCAGATGCGCTGTTTCATCTACAGGTCAAGTTCATTGGC 431
Qy 364 TTGAACCTTCTCCCAATGGACCTGTTATGAGAGAAAGACACAGGGGTGGAAACCCCAAC 423
Db 432 GTGAACCTTCTCCGTGAGACCTGTTATGCAAAAGAAAGACAATGGGCTGGGAAGCCAGC 491

Qy 424 ACTGAGCGTCTCTTTTGACGAGATGGAATGCTGATAGGAAACAACCTTTTATGGCTCTGAAG 483
Db 492 ACTGAGCGTTTGTATCTCGTGATGGCGTGTGAAAGGAGAGATTCTAATGAGCTCTGAAG 551
Qy 484 TTAGAAGAGGTGTCTACTATTTGTGTGAATTCAAATCTACTTCAAGGCGAAGGAGCCT 543
Db 552 CTGAAGACGCGTGTCTACTTACCTAGTTGAATTCAAAAGTATTTACATGGCAAGAAGCCT 611
Qy 544 GTGAAGATGCCAGGTTATCACTATGTTGACCGCAAACTGGATGTAACCAATCAACAAG 603
Db 612 GTGACGTACCAAGGTTACTATGTTGACTCCAAACTGGATATACAAGCCACAAGAA 671
Qy 604 GATTACACTTCCGTTGAGCAGCGTGAATTTCCATTGACGCAAAACCTTTTGGTC 657
Db 672 GACTATACATCGTTGAGCAGTATGAAAGAACCGGAGGAGCCACCATTCTGTC 725

Search completed: August 14, 2004, 11:00:51
Job time : 466.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 06:12:47 ; Search time 24.1414 Seconds
(without alignments)
493.990 Million cell updates/sec

Title: US-09-890-463-3
Perfect score: 1268
Sequence: 1 SVIAKQTYKYVMSGTVNGH.....SIARKPLVACCFRVRKSRHK 231

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/ECTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	482.5	38.1	238	3	US-09-277-716-16 Sequence 16, Appl
2	482.5	38.1	238	4	US-09-609-161B-16 Sequence 16, Appl
3	482.5	38.1	238	4	US-09-626-581D-65 Sequence 65, Appl
4	482.5	38.1	238	4	US-09-415-765B-65 Sequence 65, Appl
5	482.5	38.1	238	4	US-09-626-580C-65 Sequence 65, Appl
6	473.5	37.3	238	3	US-09-277-716-32 Sequence 32, Appl
7	473.5	37.3	238	4	US-09-609-161B-32 Sequence 32, Appl
8	471.5	37.2	238	4	US-09-839-650-3 Sequence 3, Appl
9	197.5	15.6	238	4	US-09-472-065A-5 Sequence 12, Appl
10	192.5	15.2	238	4	US-09-479-645A-12 Sequence 12, Appl
11	191.5	15.1	238	4	US-09-023-946B-28 Sequence 28, Appl
12	189.5	14.9	238	4	US-09-023-946B-23 Sequence 23, Appl
13	188.5	14.9	238	4	US-09-023-946B-36 Sequence 36, Appl
14	188.5	14.9	240	4	US-09-129-192C-49 Sequence 49, Appl
15	188	14.8	238	4	US-09-023-946B-22 Sequence 22, Appl
16	187.5	14.8	238	1	US-08-337-915A-2 Sequence 2, Appl
17	187.5	14.8	238	1	US-08-753-143-2 Sequence 2, Appl
18	187.5	14.8	238	2	US-08-679-865-2 Sequence 2, Appl
19	187.5	14.8	238	2	US-08-680-876-2 Sequence 2, Appl
20	187.5	14.8	238	2	US-08-792-553-2 Sequence 2, Appl
21	187.5	14.8	238	3	US-08-893-327-16 Sequence 16, Appl
22	187.5	14.8	238	3	US-08-753-144-2 Sequence 2, Appl
23	187.5	14.8	238	3	US-09-094-359-2 Sequence 2, Appl
24	187.5	14.8	238	3	US-09-172-063-2 Sequence 2, Appl
25	187.5	14.8	238	3	US-09-121-539-1 Sequence 1, Appl
26	187.5	14.8	238	3	US-09-263-975-2 Sequence 2, Appl
27	187.5	14.8	238	4	US-08-727-452-2 Sequence 2, Appl

28	187.5	14.8	238	4	US-09-418-785-1 Sequence 1, Appl
29	187.5	14.8	238	4	US-09-214-909-2 Sequence 2, Appl
30	187.5	14.8	238	4	US-09-479-645A-10 Sequence 10, Appl
31	187.5	14.8	238	4	US-09-479-645A-159 Sequence 159, Appl
32	187.5	14.8	238	4	US-09-129-192C-2 Sequence 2, Appl
33	187.5	14.8	238	4	US-09-129-192C-74 Sequence 74, Appl
34	187.5	14.8	238	4	US-09-602-641-2 Sequence 2, Appl
35	187.5	14.8	238	4	US-09-704-463-2 Sequence 2, Appl
36	187.5	14.8	238	4	US-09-472-065A-4 Sequence 4, Appl
37	187.5	14.8	238	4	US-09-472-065A-6 Sequence 6, Appl
38	187.5	14.8	238	4	US-09-023-946B-29 Sequence 29, Appl
39	187.5	14.8	238	4	US-09-920-922-4 Sequence 4, Appl
40	187.5	14.8	238	5	PCT-US95-14692-2 Sequence 2, Appl
41	187.5	14.8	239	3	US-08-646-538-2 Sequence 2, Appl
42	187.5	14.8	239	3	US-09-094-359-4 Sequence 4, Appl
43	187.5	14.8	239	3	US-09-172-063-3 Sequence 3, Appl
44	187.5	14.8	239	3	US-09-503-222-2 Sequence 2, Appl
45	187.5	14.8	239	4	US-09-513-783A-46 Sequence 46, Appl

ALIGNMENTS

RESULT 1
US-09-277-716-16
; Sequence 16, Application US/09277716A
; Patent No. 6232107

; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: PROLUME, LTD.
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE
; CURRENT APPLICATION NUMBER: US/09/277,716A
; CURRENT FILING DATE: 1999-03-26
; EARLIER APPLICATION NUMBER: 60/102,939
; EARLIER FILING DATE: 1998-10-01
; EARLIER APPLICATION NUMBER: 60/089,367
; EARLIER FILING DATE: 1998-06-15
; EARLIER APPLICATION NUMBER: 60/079,624
; EARLIER FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16

; LENGTH: 238

; TYPE: PRT

; ORGANISM: Renilla mulleri

; FEATURE:

; OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)

US-09-277-716-16

Query Match 38.1%; Score 482.5; DB 3; Length 238;
Best Local Similarity 44.0%; Pred. No. 2.2e-45;
Matches 96; Conservative 41; Mismatches 74; Indels 7; Gaps 3;

QY	7	MTYKVMGTVNGHYFEVGGDKGPKYGEQTVRLAVTKGGLPFAWDILSPQCOYGSP 66
DB	15	MSYKVNLSGIYNNHVFTWEGCGKNILFGNQLVQIRVTKGAPLPAFLVSPAFQYGNRT 74
QY	67	FTKYPEDIPDYVKQSPGGRYTWERIMNFEDGAVCTVSDSSIQGNCFIYHVKFSGLNFPP 126
DB	75	FTKYPNDISDYFIQSPFAGFYERTLRYEDGGVGLVIRSDINLIEDKFVYRVYKGSNFPD 134
QY	127	NGPVNKKYQGWEPNTERLFPARDGMLIGNNPMALKEGGHYLCFEKFKYKARKVK-MP 185
DB	135	DGPVNQKTLTGIEPSFAMYNMNGVLGVILVYKNSGYKYSCHMKTKLMSKGVVKGFPP 194
QY	186	GYHYVDKLDVTHNKNKYTSVEQREISAR-----KPL 218
DB	195	SYHFQHRLEKT-YVEDGGFVEQHETATAQMTSICKPL 231

RESULT 2

US-09-609-161B-16

```
; Sequence 16, Application US/09609161B
; Patent No. 6436682
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: PROLUME, LTD.
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE LUC
; TITLE OF INVENTION: AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HI
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
; FILE REFERENCE: 24729-121B
; CURRENT APPLICATION NUMBER: US/09/609,161B
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/277,716
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/102,939
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/089,367
; PRIOR FILING DATE: 1998-06-15
; PRIOR APPLICATION NUMBER: 60/079,624
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 16
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla mulleri
; FEATURE:
; OTHER INFORMATION: Renilla mulleri Green Fluorescent Protein (GFP)
US-09-609-161B-16

Query Match      38.1%; Score 482.5; DB 4; Length 238;
Best Local Similarity 44.0%; Pred. No. 2.2e-45;
Matches 96; Conservative 41; Mismatches 74; Indels 7; Gaps 3;

QY 7 MTKYVMSGTVNGHYFVEVGDKGKPYEGEQTIVRLAVTKGGPLPFAWDILSPQCYGSSIP 66
DB 15 MSYKVNLEGIVNNHVFVTMEGGCGKNILFNGNLQVIRVTKGAPLPFAFDIVSPAFQYGNRT 74
QY 67 FTKYPEDIPDYVKOSFPGRYTWERIMNPDGAVCTVSDSIQGNCFIYHKFSGLNPPP 126
DB 75 FTKYPNDISDYFIQSFPAGFYERTLYEDGGLVEIRSDINLIEDKFVYRVEYKGSNFPD 134
QY 127 NGPVWQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGHYLCEFFKSTYKARKPVK-MP 185
DB 135 DGPVWQKKTILGIEPSFEAMYNNGVILVYKLSGKYYSCHMTKLSKGVVKEFP 194
QY 186 GYHVVDKLDVNTNKNKYTSVEQREISAR-----KPL 218
DB 195 SYHFIQHRLEKT-YVEDGGFVEQHEHTAIAQWTSIGKPL 231

RESULT 3
US-09-626-581D-65
; Sequence 65, Application US/09626581D
; Patent No. 6548249
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
; TITLE OF INVENTION: Libraries
; FILE REFERENCE: A-66900-3/RMS
; CURRENT APPLICATION NUMBER: US/09/626,581D
; CURRENT FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 09/169,015
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 09/415,765
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 65
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla mulleri
US-09-626-581D-65

Query Match      38.1%; Score 482.5; DB 4; Length 238;
Best Local Similarity 44.0%; Pred. No. 2.2e-45;
Matches 96; Conservative 41; Mismatches 74; Indels 7; Gaps 3;

QY 7 MTKYVMSGTVNGHYFVEVGDKGKPYEGEQTIVRLAVTKGGPLPFAWDILSPQCYGSSIP 66
DB 15 MSYKVNLEGIVNNHVFVTMEGGCGKNILFNGNLQVIRVTKGAPLPFAFDIVSPAFQYGNRT 74
QY 67 FTKYPEDIPDYVKOSFPGRYTWERIMNPDGAVCTVSDSIQGNCFIYHKFSGLNPPP 126
DB 75 FTKYPNDISDYFIQSFPAGFYERTLYEDGGLVEIRSDINLIEDKFVYRVEYKGSNFPD 134
QY 127 NGPVWQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGHYLCEFFKSTYKARKPVK-MP 185
DB 135 DGPVWQKKTILGIEPSFEAMYNNGVILVYKLSGKYYSCHMTKLSKGVVKEFP 194
QY 186 GYHVVDKLDVNTNKNKYTSVEQREISAR-----KPL 218
DB 195 SYHFIQHRLEKT-YVEDGGFVEQHEHTAIAQWTSIGKPL 231

RESULT 3
US-09-626-581D-65
; Sequence 65, Application US/09626581D
; Patent No. 6548249
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
; TITLE OF INVENTION: Libraries
; FILE REFERENCE: A-66900-3/RMS
; CURRENT APPLICATION NUMBER: US/09/626,581D
; CURRENT FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: 09/169,015
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 09/415,765
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 65
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla mulleri
US-09-626-581D-65
```

```
Query Match      38.1%; Score 482.5; DB 4; Length 238;
Best Local Similarity 44.0%; Pred. No. 2.2e-45;
Matches 96; Conservative 41; Mismatches 74; Indels 7; Gaps 3;

QY 7 MTKYVMSGTVNGHYFVEVGDKGKPYEGEQTIVRLAVTKGGPLPFAWDILSPQCYGSSIP 66
DB 15 MSYKVNLEGIVNNHVFVTMEGGCGKNILFNGNLQVIRVTKGAPLPFAFDIVSPAFQYGNRT 74
QY 67 FTKYPEDIPDYVKOSFPGRYTWERIMNPDGAVCTVSDSIQGNCFIYHKFSGLNPPP 126
DB 75 FTKYPNDISDYFIQSFPAGFYERTLYEDGGLVEIRSDINLIEDKFVYRVEYKGSNFPD 134
QY 127 NGPVWQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGHYLCEFFKSTYKARKPVK-MP 185
DB 135 DGPVWQKKTILGIEPSFEAMYNNGVILVYKLSGKYYSCHMTKLSKGVVKEFP 194
QY 186 GYHVVDKLDVNTNKNKYTSVEQREISAR-----KPL 218
DB 195 SYHFIQHRLEKT-YVEDGGFVEQHEHTAIAQWTSIGKPL 231

RESULT 4
US-09-415-765B-65
; Sequence 65, Application US/09415765B
; Patent No. 6548632
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; TITLE OF INVENTION: Fusions of Scaffold Proteins with Random Peptide
; TITLE OF INVENTION: Libraries
; FILE REFERENCE: A66900-1/RMS/AMS
; CURRENT APPLICATION NUMBER: US/09/415,765B
; CURRENT FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: 09/169,015
; PRIOR FILING DATE: 1998-10-08
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 65
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Renilla mulleri
US-09-415-765B-65

Query Match      38.1%; Score 482.5; DB 4; Length 238;
Best Local Similarity 44.0%; Pred. No. 2.2e-45;
Matches 96; Conservative 41; Mismatches 74; Indels 7; Gaps 3;

QY 7 MTKYVMSGTVNGHYFVEVGDKGKPYEGEQTIVRLAVTKGGPLPFAWDILSPQCYGSSIP 66
DB 15 MSYKVNLEGIVNNHVFVTMEGGCGKNILFNGNLQVIRVTKGAPLPFAFDIVSPAFQYGNRT 74
QY 67 FTKYPEDIPDYVKOSFPGRYTWERIMNPDGAVCTVSDSIQGNCFIYHKFSGLNPPP 126
DB 75 FTKYPNDISDYFIQSFPAGFYERTLYEDGGLVEIRSDINLIEDKFVYRVEYKGSNFPD 134
QY 127 NGPVWQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGHYLCEFFKSTYKARKPVK-MP 185
DB 135 DGPVWQKKTILGIEPSFEAMYNNGVILVYKLSGKYYSCHMTKLSKGVVKEFP 194
QY 186 GYHVVDKLDVNTNKNKYTSVEQREISAR-----KPL 218
DB 195 SYHFIQHRLEKT-YVEDGGFVEQHEHTAIAQWTSIGKPL 231

RESULT 5
US-09-626-580C-65
; Sequence 65, Application US/09626580C
; Patent No. 6562617
; GENERAL INFORMATION:
; APPLICANT: Anderson, David
; APPLICANT: Bogenberger, Jakob M.
; APPLICANT: Peale, Beau R.
; TITLE OF INVENTION: FUSIONS OF SCAFFOLD PROTEINS WITH RANDOM PEPTIDE LIBRARIES
```

FILE REFERENCE: A-66900-2/RMS/AMS
CURRENT APPLICATION NUMBER: US/09/626,580C
CURRENT FILING DATE: 2000-07-27
PRIOR APPLICATION NUMBER: US 09/415,765
PRIOR FILING DATE: 1999-10-08
PRIOR APPLICATION NUMBER: US 09/169,015
PRIOR FILING DATE: 1998-10-08
NUMBER OF SEQ ID NOS: 65
SOFTWARE: PatentIn version 3.1
SEQ ID NO 65
LENGTH: 238
TYPE: PRT
ORGANISM: Renilla muelleri
US-09-626-580C-65

Query Match 38.1%; Score 482.5; DB 4; Length 238;
Best Local Similarity 44.0%; Pred. No. 2.2e-44;
Matches 96; Conservative 41; Mismatches 74; Indels 7; Gaps 3;
QY 7 MTKYVMSGTGNGHYFEVEGDKGKPYEGEQTIVRLAVTKGGPLPAWDILSPQCYGSLP 66
DB 15 MSYKNLEGI VNNHVFTEGCGKGNILFGNLQVIRVTKGAPLPFAFDIVSFAFYGNRT 74
QY 67 FTKYPEDIPDYVQSGPPGRYTWERIMNFDGAVCTVSDSSIQGNCFIYHVKFSGLNFP 126
DB 75 FTKYENDISDFIQSPAGFYVETRLRYEDGGLVEIRSDINLIEDKFVYRVEYKSNFPD 134
QY 127 NGPVMQKKTQGWENPTELFARDGMLGNFMALKEGGHYLCFEKSTYKARKPVK-MP 185
DB 135 DGPVMQKKTILGTPSFEAMYNMNGVLGVEVLIVYKLSGKYSCHMKTKLMSKGVVKEFP 194
QY 186 GYHYVDRKLDVTHNKNKYTSVEQREISAR-----KPL 218
DB 195 SYHFIQHRLEKT-YVEDGGFVEQHETAIAQMTSICKPL 231

RESULT 6
US-09-277-716-32
Sequence 32, Application US/09277716A
Patent No. 6232107
GENERAL INFORMATION:
APPLICANT: Bryan, Bruce
APPLICANT: Szent-Gyorgyi, Christopher
APPLICANT: PROLUME, LTD.
TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE
FILE REFERENCE: US/09/277,716A
CURRENT APPLICATION NUMBER: US/09-277,716A
PRIOR FILING DATE: 1999-03-26
EARLIER APPLICATION NUMBER: 60/102,939
PRIOR FILING DATE: 1998-10-01
EARLIER APPLICATION NUMBER: 60/089,367
EARLIER FILING DATE: 1998-06-15
EARLIER APPLICATION NUMBER: 60/079,624
EARLIER FILING DATE: 1998-03-27
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 32
LENGTH: 238
TYPE: PRT
ORGANISM: Ptilosarcus gurneyi
FEATURE:
OTHER INFORMATION: Ptilosarcus gurneyi Green Fluorescent Protein (GFP)
US-09-277-716-32

Query Match 37.3%; Score 473.5; DB 3; Length 238;
Best Local Similarity 44.0%; Pred. No. 2.2e-44;
Matches 96; Conservative 42; Mismatches 73; Indels 7; Gaps 3;
QY 7 MTKYVMSGTGNGHYFEVEGDKGKPYEGEQTIVRLAVTKGGPLPAWDILSPQCYGSLP 66
DB 15 MSKASVEGIVNNHVFSGEGKGNVLFQNLQVIRVTKGAPLPFAFDIVSFAFYGNRT 74
QY 67 FTKYPEDIPDYVQSGPPGRYTWERIMNFDGAVCTVSDSSIQGNCFIYHVKFSGLNFP 126

Db 75 FTKYPDDIADYFVQSPAGFFYERNLRFEDGALVDIRSDISLEDDKFKHYKVEYRGNGFPS 134
QY 127 NGPVMQKKTQGWENPTELFARDGMLGNFMALKEGGHYLCFEKSTYKARKPVK-MP 185
DB 135 NGPVMQKAILGMEFSEFVVMNNGVLGVEVLIVYKLSGNTYSCMKTFYRSKGVKGFPP 194
QY 186 GYHYVDRKLDVTHNKNKYTSVEQREISAR-----KPL 218
DB 195 EYHFIHRLLEKT-YVEGGSFVEQHETAIAQMTSICKPL 231

RESULT 7
US-09-609-161B-32
Sequence 32, Application US/09609161B
Patent No. 6436882
GENERAL INFORMATION:
APPLICANT: Bryan, Bruce
APPLICANT: Szent-Gyorgyi, Christopher
APPLICANT: PROLUME, LTD.
TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE LUCI
TITLE OF INVENTION: AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH
TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
FILE REFERENCE: 24729-121B
CURRENT APPLICATION NUMBER: US/09/609,161B
CURRENT FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/277,716
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/102,939
PRIOR FILING DATE: 1998-10-01
PRIOR APPLICATION NUMBER: 60/089,367
PRIOR FILING DATE: 1998-06-15
PRIOR APPLICATION NUMBER: 60/079,624
PRIOR FILING DATE: 1998-03-27
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 32
LENGTH: 238
TYPE: PRT
ORGANISM: Ptilosarcus gurneyi
FEATURE:
OTHER INFORMATION: Ptilosarcus gurneyi Green Fluorescent Protein (GFP)
US-09-609-161B-32

Query Match 37.3%; Score 473.5; DB 4; Length 238;
Best Local Similarity 44.0%; Pred. No. 2.2e-44;
Matches 96; Conservative 42; Mismatches 73; Indels 7; Gaps 3;
QY 7 MTKYVMSGTGNGHYFEVEGDKGKPYEGEQTIVRLAVTKGGPLPAWDILSPQCYGSLP 66
DB 15 MSKASVEGIVNNHVFSGEGKGNVLFQNLQVIRVTKGAPLPFAFDIVSFAFYGNRT 74
QY 67 FTKYPEDIPDYVQSGPPGRYTWERIMNFDGAVCTVSDSSIQGNCFIYHVKFSGLNFP 126
DB 75 FTKYPDDIADYFVQSPAGFFYERNLRFEDGALVDIRSDISLEDDKFKHYKVEYRGNGFPS 134
QY 127 NGPVMQKKTQGWENPTELFARDGMLGNFMALKEGGHYLCFEKSTYKARKPVK-MP 185
DB 135 NGPVMQKAILGMEFSEFVVMNNGVLGVEVLIVYKLSGNTYSCMKTFYRSKGVKGFPP 194
QY 186 GYHYVDRKLDVTHNKNKYTSVEQREISAR-----KPL 218
DB 195 EYHFIHRLLEKT-YVEGGSFVEQHETAIAQMTSICKPL 231

RESULT 8
US-09-839-650-3
Sequence 3, Application US/09839650
Patent No. 6645761
GENERAL INFORMATION:
APPLICANT: Stratagene
TITLE OF INVENTION: Humanized Polynucleotide Sequence Encoding Renilla Mulleri Green
Patent No. 6645761
TITLE OF INVENTION: Fluorescent Protein

Db 192 PVLLPDHXL 201

Search completed: August 12, 2004, 06:21:06
Job time : 25.1414 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 06:12:47 ; Search time 21.6701 Seconds
(without alignments)
1043.144 Million cell updates/sec

Title: US-09-890-463-4

Perfect score: 1287

Sequence: 1 SVIAKQMTYKYVMSGTVNGH.....KPVVACFFRVKSRHKYAVA 235

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	185.5	14.4	238	1 JQ1514	green-fluorescent
2	97	7.5	583	2 S84909	probable membrane
3	86.5	6.7	368	2 A81289	hypothetical prote
4	86	6.7	770	2 S00643	anthranilate synth
5	84	6.5	725	1 IQMSNG	neural cell adhesi
6	84	6.5	787	2 E82323	organic solvent to
7	83	6.4	260	2 B70419	hypothetical prote
8	83	6.4	752	2 S23818	hypothetical prote
9	82.5	6.4	600	2 T34757	probable oligopept
10	80.5	6.3	390	2 D64903	probable arylsulfa
11	80.5	6.3	3461	2 S58870	resilin precursor -
12	79.5	6.2	876	2 A89944	alanyl-tRNA synthet
13	79	6.1	403	2 T51828	probable photost
14	79	6.1	687	1 B32382	ubiquinol-cytochr
15	79	6.1	898	1 DJBPT4	DNA-directed DNA p
16	79	6.1	5175	2 T20992	hypothetical prote
17	79	6.1	5198	2 T43290	hemocytin precurs
18	78.5	6.1	363	2 S30149	cysteine proteinas
19	78.5	6.1	390	2 F90891	probable enzyme li
20	78.5	6.1	390	2 C85726	probable enzyme Z2
21	78.5	6.1	870	2 T47454	lipoxigenase ATLOX
22	78.5	6.1	896	2 JQ2391	lipoxigenase (EC 1
23	78	6.1	334	2 D95982	hypothetical expor
24	78	6.1	380	2 T00683	hypothetical prote
25	78	6.1	6642	2 T23757	protein UNC-89 - C
26	77.5	6.0	207	2 A48608	E1 glycoprotein -
27	77.5	6.0	266	1 CDPJ13	chlorophyll a/b-bi
28	77.5	6.0	297	2 E71425	hypothetical prote
29	77.5	6.0	333	2 B64380	hypothetical prote

ALIGNMENTS

RESULT 1

JQ1514

green-fluorescent protein [validated] - hydromedusa (Aequorea victoria)

C:Species: Aequorea victoria

C:Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #text change 23-Mar-2001

C:Accession: J50692; JQ1514; F00335; S48693; S51330; S51331

R:Fraser, D.C.; Eckenrode, V.K.; Ward, W.W.; Prendergast, F.G.; Cormier, M.J.

Gene 111, 229-233, 1992

A:Title: Primary structure of the Aequorea victoria green-fluorescent protein.

A:Reference number: JQ1514; MUID:92175527; PMID:1347277

A:Accession: J50692

A:Molecule type: DNA

A:Residues: 1-107, 'S', 109-238 <PRAL>

A:Cross-references: GB:M62654; NID:g155662; PIDN:AAA27722.1; PID:g155663

A:Accession: JQ1514

A:Molecule type: mRNA

A:Residues: 1-99, 'P', 101-140, 'L', 142-218, 'V', 220-238 <PRA2>

A:Cross-references: GB:M62653; NID:g155660; PIDN:AAA27721.1; PID:g155661

A:Accession: F00335

A:Molecule type: Protein

A:Residues: 46-64;74-122;132-151;154-183;185-200 <PRA3>

R:Inouye, S.; Tsuji, F.I.

PNBS Lett. 351, 211-214, 1994

A:Title: Evidence for redox forms of the Aequorea green fluorescent protein.

A:Reference number: S48693; MUID:94364470; PMID:8082767

A:Accession: S48693

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-24, 'Q', 26-156, 'P', 158-171, 'K', 173-238 <INO>

A:Cross-references: GB:L29345; NID:g606383; PIDN:AAA58246.1; PID:g606384

R:Watkins, J.N.; Campbell, A.K.

submitted to the EMBL Data Library, January 1995

A:Reference number: S51330

A:Accession: S51330

A:Molecule type: mRNA

A:Residues: 1-13, 'V', 15-24, 'Q', 26-44, 'N', 46-153, 'G', 155-156, 'P', 158-171, 'K', 173-227, 'R', 2

A:Cross-references: EMBL:X83960; NID:g634008; PIDN:CAA58789.1; PID:g634009

A:Experimental source: clone gfp1

A:Accession: S51331

A:Molecule type: mRNA

A:Residues: 1-24, 'Q', 26-29, 'R', 31-83, 'L', 85-153, 'G', 155-156, 'P', 158-171, 'K', 173-208, 'Q', 2

A:Cross-references: EMBL:X83960; NID:g634010; PIDN:CAA58790.1; PID:g634011

A:Experimental source: clone gfp2

R:Yang, F.; Moss, L.G.; Phillips Jr., G.N.

submitted to the Brookhaven Protein Data Bank, August 1996

A:Reference number: A65692; PDB:1GFL

A:Contents: annotation: X-ray crystallography, 1.9 angstroms, residues 'A', 2-79, 'R', 81-95

A>Note: engineered sequence based on JQ1514, cloned and expressed in Escherichia coli

R:Yang, F.; Moss, L.G.; Phillips Jr., G.N.

Nat. Biotechnol. 14, 1246-1251, 1996

A:Title: The molecular structure of green fluorescent protein.

A:Reference number: A58953; MUID:98294543; PMID:9631087

A;Contents: annotation; X-ray crystallography, 1.9 angstroms
C;Comment: This protein is excited by the photoprotein aequorin (see PIR:AQJFNV) emitting
C;Comment: The chromophore of this protein is formed by modification of Ser-dehydro-Tyr-
C;Genetics: C;
A;Gene: GFP
A;Introns: 69/3; 167/3
C;Superfamily: green-fluorescent protein
C;Keywords: chromoprotein; luminescence
F;65-67/Cross-link: 5-imidazolone (Ser-Gly) #status experimental
F;66/Modified site: dehydroxyrosine (Tyr) #status experimental

Query Match 14.4%; Score 185.5; DB 1; Length 238;
Best Local Similarity 24.8%; Pred. No. 4.9e-09;
Matches 51; Conservative 47; Mismatches 91; Indels 17; Gaps 8;

Qy 11 VYMSGTVNGHYFEVGGDKGPKYEGEQTVRLAVTKGGPLFAWDILSPQCOYGSIPFTKY 70
Db 16 VELGDVNGHKFSVSGEGGATYKGLTKFICTT-GKLPVWPVTLVTFYSGVQCFSRY 74
Qy 71 PEDIP--DYVKQSPFGRTYMERIMNFEDGAVCTVSNDSIQGNCFIYHKFSGLNFPNG 128
Db 75 PDHMKQHDFFKSAPEGVQVQRTIFYKDDGNYKTRAEVKFEGDTLVNRIELKGDIFKEDG 134
Qy 129 PVMQKKTQGWEPNTERLP-----ARDGMLIGNNF-MALKLEGSGHYLCEP--KSTYKAKK 180
Db 135 NILGHRME-YNYNSHNVYIMADKQNGIKV--NFKIRHNIEDGSGVQLADHYQQNTPIGDG 191
Qy 181 PVKMPGVHYVDRKLDVT---NHNKDY 203
Db 192 PVLLPDNHYLSTQALSNDKPEKRDH 217

RESULT 2
S64909
probable membrane protein YLR077w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein L2349
C;Species: Saccharomyces cerevisiae
C;Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 19-Apr-2002
C;Accession: S64909
R;Pohl, T.M.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64899
A;Accession: S64909
A;Molecule type: DNA
A;Residues: 1-583 <POH>
A;Cross-references: EMBL:Z73249; NID:g1360436; PID:e245794; PID:g1360437; GSPDB:GN00012;
A;Experimental source: strain S288C
C;Genetics:
A;Gene: MIPS:YLR077w
A;Cross-references: SGD:S0004067
A;Map position: 12R
C;Superfamily: Saccharomyces cerevisiae probable membrane protein YLR077w
C;Keywords: transmembrane protein
F;81-97/Domain: transmembrane #status predicted <TM>

Query Match 7.5%; Score 97; DB 2; Length 583;
Best Local Similarity 20.6%; Pred. No. 1.1;
Matches 55; Conservative 33; Mismatches 89; Indels 90; Gaps 13;

Qy 18 NGHYFEVGGDKGPKYEGEQTVR-----LAVTKGGPL-PEAWDILSPQCOY 62
Db 349 NNQLFEIELLNKFK-HGEDVVRKREIKTKACGSYHTALDKTGEIYAFGMN-----RF 401
Qy 63 G-----SIPFTKYPEDIPDYVKQSPFGRTYMERI----- 91
Db 402 GQLALPISYNLEYVSFPRSVTHAFKPHFGPMGNWKKCVDIHDDDETSFVTRIKPGSTDHH 461
Qy 92 -MNFEDGAVCTVSNDSIQGNCFIYHKFSGLNFPNGPVNQKKTQGWEPNTERLPA--- 147
Db 462 YFAPGNLFLGELGNLTFFNQCDDPIKTKSD-----DKKLWNKSGSHCVFTETE 510
Qy 148 --RDGMLIGNNFALKLEGSGH-YLCEFKSTYKAKPKVKM-----PGHYVDRKLDVTNH 199

Db 511 QENEVIAGNN-----DHGQLGIGKTKMKCAKPMNIEVLKPGQDTTD--LDSIYN 559
Qy 200 NKDYTSVEQCEISIAKRPVWACRRFRV 226
Db 560 SKLHLKKEQRVVTNGNK----SCLYWRV 583

RESULT 3
A81289
hypothetical protein Cj1433c [imported] - Campylobacter jejuni (strain NCTC 11168)
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C;Accession: A81289
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chilling
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp
A;Reference number: A81250; MUID:20150912; PMID:10698204
A;Accession: A81289
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-368 <PAR>
A;Cross-references: GB:AL139078; GB:AL111168; NID:9698723; PIDN:CAB73857.1; PID:g969886;
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: Cj1433c
C;Superfamily: Campylobacter jejuni hypothetical protein Cj1433c

Query Match 6.7%; Score 86.5; DB 2; Length 368;
Best Local Similarity 21.9%; Pred. No. 5.1;
Matches 46; Conservative 22; Mismatches 51; Indels 91; Gaps 12;

Qy 48 PLPFAWDILSPQCOYQGSIPFTKYPEDIPDYVKQSPFGRTYMERIMNFEDGAVCTVSNDS 106
Db 227 PIKINWLIKDKDOY-----DIP-----LIFFNNGELEKTSWKF 261
Qy 107 SI--QGNCFYHYKFSGLNFPNGPVNQKKTQGWEPNTERLFARDGMLIGNFMA----L 160
Db 262 SLDPGSGNCDYH---SFTNGSMANHCYQFK-----DGKLTCTTFFAHVQHF 304
Qy 161 KLEGGHY-LCEFK--STYKAKPKVPKMPGVHYVDRKLDVTNHKDYTSVEQCEISIAKRP 217
Db 305 NKKYGNHFVCECFDFDIYKA-----KDYQEI-----LFFLSKP 338
Qy 218 WVACRRFRV-----KSRHKYAV 234
Db 339 IPFCRYCKVQSWAEIGKWRSSNKTKEYLI 368

RESULT 4
S00643
anthranilate synthase multifunctional protein - Aspergillus niger
N;Alternate names: gene trpG-trpC-trpF protein; glutamine amidotransferase
N;Contains: anthranilate synthase (EC 4.1.3.27); indole-3-glycerol-phosphate synthase (EC
C;Species: Aspergillus niger
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 31-Mar-2000
C;Accession: S00643; A23979; B23979
R;Kos, T.; Kuijvenhoven, A.; Hessing, H.G.M.; Pouwels, P.H.; van den Hondel, C.A.M.J.J.
Curr. Genet. 13, 137-144, 1988
A;Title: Nucleotide sequence of the Aspergillus niger trpC gene: structural relationship
A;Reference number: S00643; MUID:88223483; PMID:2836085
A;Accession: S00643
A;Molecule type: DNA
A;Residues: 1-770 <KOS>
A;Cross-references: EMBL:X07071; NID:g2420; PIDN:CAA30107.1; PID:g2421
R;Kos, A.; Kuijvenhoven, J.; Wernars, K.; Bos, C.J.; van den Broek, H.W.J.; Pouwels, P.H.
Gene 39, 231-238, 1985
A;Reference number: A91539; MUID:86137391; PMID:2936650
A;Accession: A23979
A;Molecule type: DNA
A;Residues: 1-69 <KO2>
A;Cross-references: GB:ML4404; NID:g166536; PIDN:AAA32709.1; PID:g166537
A;Accession: B23979

```

Query Match      6.5%; Score 84; DB 1; Length 725;
Best Local Similarity 22.9%; Pred. No. 20;
Matches 54; Conservative 38; Mismatches 82; Indels

QY      10  KYVMSTVNGHYFEVGGDGKG--PYE-----GEQIVR-LAVTKGGP
Db       455  KII--NTPSASYLEVTPDSENGFNVCYAVNRIGESLEFLTVQADPTF
QY      59  ----CQCY-----GSIPRTKYPEIDPYDKSFPFGRYTWIRIMNFEDG
Db       513  SSTAQVQFPEPEATGVPILKYAEKWSLGBESW--HFTWYD-----
QY     109  QGNCFI-----YHVESGLNFPNGPVMQ---KKTQG-----WEPNVT
Db     559  EGIIVTIGLKPEITTSYSDRLAAALNGKLGEIMQFSEKTPQVPEUSAPKLI
QY     154  GNNFMALKEGGG---HYLCBFKSYTKAKK-VKMP-GYHYVDRKLDV
Db     619  KYNL--IKDDGSGSPTRHYLVKYRALASEWKPEIRLPSGSHHV--MKKSL

```

```

RESULT 6
E82323      organic solvent tolerance protein VC0446 [imported] - Vibrio cholerae (strain N16961 sero-
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: E82323
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: E82323
A>Status: preliminary
A:Molecule type: DNA
A;Residues: 1-787 <HEI>
A;Cross-references: GB:AE004131; GB:AE003852; NID:g9654856; PIDN:AAF93619.1; GSPDB:GN0012
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0446
A;Map position: 1
C;Superfamily: organic solvent tolerance protein

Query Match          6.5%; Score 84; DB 2; Length 787;
Best Local Similarity 18.8%; Pred. No. 22;
Matches 44; Conservative 34; Mismatches 92; Indels 64; Gaps 9;

QY       6 QMTVKVMSGTVN-----CHYEVEGDGKGKPYEGQTIVRLAVTKGGPLPPA--WDIL 56
Db       388 QLSNYAPETPMKYLDLVLGVHSRSFETARGKP-----SATRVHIEPGLKIPFSNTWGNW 443

QY       57 SPQCQGSPIPTKYPEDIPDVYKQSFPGRYTWERIMNPEDGACVTYSNDSSIOGCNFYIH 116
Db       444 ITEAR---VLGYTQQDLXKTDD-----AKUEESVTRVIPETRVS----- 480

QY       117 VKFGSLNFFPNPGPVQMCKTGWEFNTERLFPA-----RDGMGLIGNFMAL----- 160

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 06:12:47 ; Search time 14.4467 Seconds
(without alignments)
847.008 Million cell updates/sec

Title: US-09-890-463-4

Perfect score: 1287

Sequence: 1 SVIAKQMTYKYVMGTVNGH.....KPVVACRFVRVKSRRHKYAVA 235

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	188.5	14.6	238	1 GFP_ABOVI	P42212 aequorea vi
2	86	6.7	770	1 TRPG_ASPNG	P05328 aspergillus
3	84	6.5	725	1 NCA2_MOUSE	P13594 mus musculus
4	84	6.5	787	1 OSTA_VIBCH	Q9kur9 vibrio chol
5	83	6.4	260	1 YD69_AQUAE	Q67381 aquifex aeo
6	81	6.3	1142	1 ENAM_HUMAN	Q9nmi1 homo sapien
7	80.5	6.3	385	1 YDEM_ECOLI	P76134 escherichia
8	80.5	6.3	3461	1 REIN_MOUSE	Q60841 mus musculus
9	79.5	6.2	876	1 SYA_STAAM	Q99tn1 staphylococ
10	79.5	6.2	876	1 SYA_STAAM	Q8hw87 staphylococ
11	79	6.1	403	1 H136_ARATH	O82660 arabidopsis
12	79	6.1	687	1 CYBC_BRAVA	P51131 bradyrhizob
13	79	6.1	898	1 DROL_EPT4	P04415 bacterioph
14	78.5	6.1	896	1 LOXC_ARATH	P38418 arabidopsis
15	78.5	6.1	3462	1 REIN_RAT	P58751 rattus norv
16	78	6.1	6632	1 UN89_CAEEL	O01761 caenorhabdi
17	77.5	6.0	266	1 CB21_PETGP	P04779 petunia sp.
18	77.5	6.0	333	1 Y642_METUA	Q58059 methanococc
19	77.5	6.0	366	1 MUTA_EUCAP	Q8k9a7 buchnera ap
20	77	6.0	726	1 RRP2_INBP9	O36432 influenza b
21	76.5	5.9	1526	1 YV46_ANASP	Q8yri1 anabaena sp
22	76	5.9	266	1 CB24_PETGP	P04782 petunia sp.
23	76	5.9	342	1 XVA_CALSA	P23556 caldocellum
24	76	5.9	444	1 YGAF_ECOLI	P37339 escherichia
25	76	5.9	725	1 RRP2_INBSI	P11136 influenza b
26	76	5.9	1162	1 VGL2_IBVB	P11223 avian infec
27	75.5	5.9	587	1 COAT_PAVL3	P36310 parvovirus
28	75	5.8	248	1 YA14_SCHPO	Q09686 schizosacch
29	75	5.8	267	1 CB22_PETGP	P04780 petunia sp.
30	75	5.8	392	1 ANIA_NEIGO	Q02419 neisseria g
31	75	5.8	905	1 HXAI_HAEIN	P44602 haemophilus
32	75	5.8	1004	1 MV10_MOUSE	P23249 mus musculus
33	74.5	5.8	513	1 YHCX_BACSU	P54608 bacillus su

ALIGNMENTS

RESULT 1				
GFP_AEQVI				
ID	GFP_AEQVI	STANDARD;	PRT;	238 AA.
AC	P42212; Q17104; Q27903;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Green fluorescent protein.			
GN	GFP.			
OS	Aequorea victoria (Jellyfish).			
OC	Eukaryota; Metazoa; Chnidaria; Hydrozoa; Hydroida; Leptomedusae;			
OC	Aequoreidae; Aequorea.			
OX	NCBI_TaxID=6100;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RX	MEDLINE=92175527; PubMed=1347277;			
RA	Prasher D.C., Eckenrode V.K., Ward W.W., Prendergast F.G., Cormier M.J.;			
RT	"Primary structure of the Aequorea victoria green-fluorescent protein.";			
RL	Gene 111:229-233 (1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94185810; PubMed=8137953;			
RA	Inouye S., Tsuji F.I.;			
RT	"Aequorea green fluorescent protein. Expression of the gene and fluorescence characteristics of the recombinant protein.";			
RL	FEBS Lett. 341:277-280 (1994).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97299832; PubMed=9154981;			
RA	Rouwendal G.J.A., Mendes O., Wolbert E.J.H., de Boer A.D.;			
RT	"Enhanced expression in tobacco of the gene encoding green fluorescent protein by modification of its codon usage.";			
RL	Plant Mol. Biol. 33:989-999 (1997).			
RN	[4]			
RP	CHROMOPHORE.			
RX	MEDLINE=93192221; PubMed=8448132;			
RA	Cody C.W., Prasher D.C., Westler W.M., Prendergast F.G., Ward W.W.;			
RT	"Chemical structure of the hexapeptide chromophore of the Aequorea green-fluorescent protein.";			
RL	Biochemistry 32:1212-1218 (1993).			
RN	[5]			
RP	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).			
RX	MEDLINE=96355665; PubMed=8703075;			
RA	Orme M., Cubitt A.B., Kallio K., Gross L.A., Tsien R.Y., Remington S.J.;			
RT	"Crystal structure of the Aequorea victoria green fluorescent protein.";			
RL	Science 273:1392-1395 (1996).			
RN	[6]			
RP	X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).			
RX	MEDLINE=98294543; PubMed=9631087;			
RA	Yang F., Moss L.G., Phillips G.N. Jr.;			
RT	"The molecular structure of green fluorescent protein.";			
RL	Nat. Biotechnol. 14:1246-1251 (1996).			

O15020 homo sapien
P07370 lycopersico
P04783 petunia sp.
Q58994 methanococ
P55631 rhizobium s
P32606 saccharomyc
P13851 sinapis alb
Q63504 rattus norv
P46200 bos taurus
O97939 sus scrofa
P39538 saccharomyc
P43316 humicola in

34 74.5 5.8 2390 1 SPCP_HUMAN
35 74 5.7 265 1 CB2B_LYCES
36 74 5.7 267 1 CB25_PETSP
37 74 5.7 277 1 YF99_METUA
38 74 5.7 398 1 Y4QJ_RHISN
39 74 5.7 800 1 PT27_YEAST
40 73.5 5.7 266 1 CB21_SINAL
41 73.5 5.7 578 1 NRD2_RAT
42 73.5 5.7 640 1 MYB_BOVIN
43 73.5 5.7 1142 1 ENAM_PIG
44 73.5 5.7 1254 1 UBPC_YEAST
45 73 5.7 213 1 GUN5_HUMIN

RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF MUTANT WITH YELLOW EMISSION.
 RP MEDLINE=98455509; PubMed=9782051;
 RA Wachter R.M., Elsiger M.A., Kallio K., Hanson G.T., Remington S.J.;
 RT "Structural basis of spectral shifts in the yellow-emission variants
 RL of green fluorescent protein.";
 RL Structure 6:1267-1277(1998).
 [8]
 RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RP MEDLINE=99238303; PubMed=10220315;
 RA Elsiger M.A., Wachter R.M., Hanson G.T., Kallio K., Remington S.J.;
 RT "Structural and spectral response of green fluorescent protein
 RL variants to changes in pH.";
 RL Biochemistry 38:5296-5301(1999).
 CC -!- FUNCTION: Energy-transfer acceptor. Its role is to transduce the
 CC blue chemiluminescence of the protein aequorin into green
 CC fluorescent light by energy transfer. Fluoresces in vivo upon
 CC receiving energy from the Ca(2+)-activated photoprotein aequorin.
 CC Absorbs light maximally at 395 nm and exhibits a smaller
 CC absorbance peak at 470 nm. The fluorescence emission spectrum
 CC peaks at 509 nm with a shoulder at 540 nm.
 CC -!- SUBUNIT: Monomer.
 CC -!- TISSUE SPECIFICITY: Photocytes.
 CC -!- PTM: Contains a covalently attached chromophore, which is composed
 CC of modified amino acid residues. The chromophore is formed upon
 CC cyclization of the residues Ser-dehydrotyr-Gly.
 CC -!- BIOTECHNOLOGY: Has become a useful and ubiquitous tool for making
 CC chimeric proteins of GFP linked to other proteins where it
 CC functions as a fluorescent protein tag. GFP tolerates N- and C-
 CC terminal fusion to a broad variety of proteins. It has been
 CC expressed in bacteria, yeast, slime mold, plants, Drosophila,
 CC zebrafish, and in mammalian cells. As a noninvasive fluorescent
 CC marker in living cells, it allows for a wide range of applications
 CC where it may function as a cell lineage tracer, reporter of gene
 CC expression, or as a measure of protein-protein interactions.
 CC -!- DATABASE: NAME=Protein Spotlight;
 CC NOTE=Issue 11 of June 2001;
 CC WWW="http://www.expasy.org/spotlight/articles/sp11011.html".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M62654; AAA27722.1; -;
 DR EMBL; M62653; AAA27721.1; -;
 DR EMBL; L29345; AAA58246.1; -;
 DR EMBL; X96418; CAA65278.1; -;
 DR PIR; JS0692; JQ1514.
 DR PDB; 1B9C; 17-NOV-00.
 DR PDB; 1BFP; 07-JUL-97.
 DR PDB; 1C4F; 14-JUN-00.
 DR PDB; 1C4F; 14-JUN-00.
 DR PDB; 1EMA; 08-NOV-96.
 DR PDB; 1EMB; 16-JUN-97.
 DR PDB; 1EMC; 20-AUG-97.
 DR PDB; 1EME; 20-AUG-97.
 DR PDB; 1EMF; 20-AUG-97.
 DR PDB; 1EMG; 12-MAY-99.
 DR PDB; 1EMH; 20-AUG-97.
 DR PDB; 1EMK; 20-AUG-97.
 DR PDB; 1EMM; 20-AUG-97.
 DR PDB; 1F09; 17-NOV-00.
 DR PDB; 1F0B; 17-NOV-00.
 DR PDB; 1GFL; 11-JAN-97.
 DR PDB; 1HCJ; 15-JAN-02.
 DR PDB; 1HUY; 04-JUL-01.
 DR PDB; 1JBY; 07-JAN-03.
 DR PDB; 1JBZ; 07-JAN-03.
 DR PDB; 1KP5; 28-AUG-02.
 DR PDB; 1KYP; 10-APR-02.

DR PDB; 1KYR; 10-APR-02.
 DR PDB; 1KYS; 10-APR-02.
 DR PDB; 1YFP; 28-OCT-98.
 DR PDB; 2EMD; 20-AUG-97.
 DR PDB; 2EMN; 20-AUG-97.
 DR PDB; 2EMO; 20-AUG-97.
 DR PDB; 2EMO; 20-AUG-97.
 DR InterPro; IPR003017; GFP_like.
 DR InterPro; IPR000786; Green_fl_protein.
 DR Pfam; PF01353; GFP; 1.
 DR PRINTS; PR01229; GFP; 1.
 DR ProDom; PD013756; Green_fl_protein; 1.
 DR Luminescence; 3D-structure.
 KW CROSSLINK 65 67 5-imidazolinone (Ser-Gly).
 FT MOD RES 66 66 2,3-DIDEHYDROTYROSINE.
 FT VARIANT 100 100 F -> Y.
 FT VARIANT 108 108 T -> S.
 FT VARIANT 141 141 L -> M.
 FT VARIANT 219 219 V -> I.
 FT CONFLICT 2 2 S -> G (IN REF. 3).
 FT CONFLICT 25 25 H -> Q (IN REF. 2).
 FT CONFLICT 80 80 Q -> R (IN REF. 3).
 FT CONFLICT 157 157 Q -> P (IN REF. 2).
 FT CONFLICT 172 172 E -> K (IN REF. 2).
 FT CONFLICT 172 172 E -> K (IN REF. 2).
 FT HELIX 4 8
 FT STRAND 12 22
 FT TURN 23 24
 FT STRAND 25 36
 FT STRAND 37 40
 FT TURN 41 48
 FT STRAND 49 50
 FT TURN 57 60
 FT HELIX 61 63
 FT TURN 69 71
 FT STRAND 73 73
 FT HELIX 76 81
 FT HELIX 83 86
 FT TURN 87 90
 FT STRAND 92 100
 FT TURN 101 102
 FT STRAND 105 115
 FT TURN 116 117
 FT STRAND 118 128
 FT TURN 132 133
 FT TURN 135 139
 FT STRAND 141 141
 FT STRAND 148 155
 FT TURN 156 159
 FT STRAND 160 171
 FT TURN 172 173
 FT STRAND 176 187
 FT STRAND 199 208
 FT TURN 211 212
 FT STRAND 217 227
 SQ SEQUENCE 238 AA; 26886 MW; EA5A6F21PFB56E05 CRC64;
 Query Match 14.6%; Score 188.5; DB 1; Length 238;
 Best Local Similarity 25.2%; Pred. No. 8.4e-10;
 Matches 52; Conservative 46; Mismatches 91; Indels 17; Gaps 8;
 Qy 11 VYMSGTVNGHYFEVGGKPKYEGEQTIVRLAVTKGGPLPEAWDLISPCQCOYGSIPETKY 70
 Db 16 VELDGDVNGHKFSVSGEGGDATYKGLTKLFICTT-GKLPVFWPTLVTFYGVQCFGRY 74
 Qy 71 PEDIP--DYVKQSPFFGRYTWERIMNFEDGAVCTVSNDSIQNCFIYHKFSGLNFPNG 128
 Db 75 PDHMKQHDFFKSPAMPEGVQERTIFFKDDGNVKTRAEVKFEGDTLVNRIELKIDGDKEDG 134
 Qy 129 PVMQKKTQGWEPNTERLF-----ARDGMLIGNNP-MALKLGGGHYLCEF--KSTYKAKK 180
 Db 135 NILGHKLE-YNYNSHNVYIMADKQNGIKV--NFKIRHNIEDGVSQVLADHYQONTPIG 191
 Qy 181 PVKMPGVHYVDRKLDVT---NHNKDY 203

"Differential splicing and alternative polyadenylation generates distinct NCAM transcripts and proteins in the mouse.";
EMBO J. 7:625-632(1988).

[4]

SEQUENCE OF 20-36.

MEDLINE=86140120; PubMed=3512556;

Rougon G., Marshak D.R.;

"Structural and immunological characterization of the amino-terminal domain of mammalian neural cell adhesion molecules.";

J. Biol. Chem. 261:3396-3401(1986).

-!- FUNCTION: This protein is a cell adhesion molecule involved in neuron-neuron adhesion, neurite fasciculation, outgrowth of neurites, etc.

-!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.

-!- ALTERNATIVE PRODUCTS:

Event=Alternative splicing; Named isoforms=3;

Name=N-CAM 120;

Isoid=PI3594-1; Sequence=Displayed;

Name=N-CAM 180;

Isoid=PI3595-1; Sequence=External;

Name=N-CAM 140;

Isoid=PI3595-2; Sequence=External;

-!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.

-!- SIMILARITY: Contains 2 fibronectin type III domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi-sib.ch/announce/> or send an email to license@ebi-sib.ch).

EMBL; Y00051; CAA68263.1; -;
EMBL; X15049; CAA33148.1; ALT_SEQ.

EMBL; X07195; CAA30173.1; -;

PIR; A29673; IJMSNG.

PDB; 3NCM; 12-MAR-97.

PDB; 3NCM; 23-JUL-99.

MGI; 97281; Ncam1.

InterPro; IPR008957; FN III-like.

InterPro; IPR003961; FNIII.

InterPro; IPR007110; IG-like.

InterPro; IPR003598; IG_c2.

Pfam; PF00041; fn3; 2.

Pfam; PF00047; ig; 5.

SMART; SMO0060; FN3; 2.

SMART; SMO0408; IGC2; 5.

PROSITE; PS00835; IG_LIKE; 5.

Cell adhesion; Glycoprotein; Repeat; Alternative splicing;

Immunoglobulin domain; Signal; Heparin-binding; GPI-anchor;

3D-structure.

SIGNAL 1 19

CHAIN 20 725

NEURAL CELL ADHESION MOLECULE 1, 120 kDa

ISOFORM.

IG-LIKE C2-TYPE 1.

IG-LIKE C2-TYPE 2.

IG-LIKE C2-TYPE 3.

IG-LIKE C2-TYPE 4.

IG-LIKE C2-TYPE 5.

FIBRONECTIN TYPE-III 1.

FIBRONECTIN TYPE-III 2.

HEPARIN-BINDING (POTENTIAL).

HEPARIN-BINDING (POTENTIAL).

PROBABLE.

PROBABLE.

PROBABLE.

PROBABLE.

PROBABLE.

PROBABLE.

PROBABLE.

PROBABLE.

PROBABLE.

PROBABLE.

PROBABLE.

PROBABLE.

PROBABLE.

PROBABLE.

PROBABLE.

FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 479 479 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 261 268 ERSRVS -> DEKHIFSD (IN REF. 2).
FT CONFLICT 273 273 V -> L (IN REF. 2).
FT CONFLICT 354 355 QD -> KT (IN REF. 2).
FT CONFLICT 549 549 T -> K (IN REF. 2).
FT CONFLICT 572 572 T -> R (IN REF. 2).
FT CONFLICT 575 575 D -> V (IN REF. 2).
FT CONFLICT 594 594 MOPSES -> SAATEF (IN REF. 2).
FT CONFLICT 600 602 PEL -> REP (IN REF. 2).
FT CONFLICT 657 657 H -> D (IN REF. 2 AND 3).
SQ SEQUENCE 725 AA; 80296 MW; C2AEB8B4461C6B2F CRC64;

Query Match 6.5%; Score 84; DB 1; Length 725;

Best Local Similarity 22.9%; Pred. No. 7.6;

Matches 54; Conservative 38; Mismatches 82; Indels 62; Gaps 16;

Qy 10 KVMGTVNGHYFEVGDGK - PYE - - - - - GEQTVR-LAVTKGGPLPFAWDILSP- 58

Db 455 KIV--NTPSASYLEVTPDSENFQVNCYNTAVNRIGQESLEFIVQADTPSPSIDRVEPY 512

Qy 59 - - - - - QCOY - - - - - GSIPFTKYPEDIPDYVKQSPGRTYWERIMNFEDGAVCTVSDSSI 108

Db 513 SSTAQVDFDEATGGVFLIKYKAWSLGEESW - HFTWYD - - - - - AKENM 558

Qy 109 QGNCFI - - - - - YHVKFSLNFPNGPVNQ - - - - - KKTQG - - - - - WEPNTERLFARDGMLI 153

Db 559 EGIVTIMGLKPEYTSYDRALANGKLGELMOPSESKTQPVPELSAPKLEQMGEDGNSI 618

Qy 154 GNNFMALKLEGG - - - - - HYLCEFKSTYKAKP - VMKP - GYHYVDRKLDVTNHNKY 203

Db 619 KVLN - - - - - IKQDDGSPIRHYLVKYRALASEWKPEIRLPSGSHV - MLKSLDWNAY 670

RESULT 4

OSTA_VIBCH STANDARD; PRT; 787 AA.

AC Q9KOR9;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DE Organic solvent tolerance protein precursor.

GN IMP OR OSTA OR VC0446.

OS Vibrio cholerae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

OC Vibrionaceae; Vibrio.

OX NCBI_TaxID=666;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=El Tor N16961 / Serotype O1;

RX MEDLINE=20406833; PubMed=10952301;

RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,

RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,

RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,

RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,

RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,

RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,

RA Fraser C.M.;

RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio

cholerae";

RL Nature 406:477-483(2000).

CC -!- FUNCTION: Determines N-hexane tolerance. Involved in outer

membrane permeability. Essential for envelope biogenesis. Could be

part of a targeting/usher system for outer membrane components (By

similarity).

CC -!- SUBCELLULAR LOCATION: Outer membrane (By similarity).

CC -!- SIMILARITY: Belongs to the imp/osta family.

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

CC


```

FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 467 467 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 534 534 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 934 934 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1040 1040 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1142 AA; 128745 MW; 77419C4375EAD6EC CRC64;

Query Match 6.3%; Score 81; DB 1; Length 1142;
Best Local Similarity 29.5%; Pred. No. 24;
Matches 23; Conservative 12; Mismatches 35; Indels 8; Gaps 3;

Qy 18 NGHVEVEGKGPYEGEGRVLAATKG-----GPIPFWDILSPQCQVGSIDFTKYPE 72
Db 493 NSYV--PRGDSRKVPNSDGTQSQNLPGKIVLGSRMPYESTNQSLKHSSYPVAVYPE 550

Qy 73 DIPDYVKQSPF-GRYTWE 89
Db 551 EIPSPAKEHPFAGRTWTD 568

RESULT 7
YDEM_ECOLI
ID_YDEM_ECOLI STANDARD; PRT; 385 AA.
AC P76134; E77755;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DB Hypothetical protein ydem.
GN YDEM OR B1497.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=K12 / MG1655;
RX MEDLINE=97428617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12.";
RA Science 277:1453-1474(1997).
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA Motomura K., Nakano S., Nakamura Y., Nishimoto H., Nishio Y.,
RA Oshima T., Saito N., Sampei G., Seki Y., Sivasubram S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horiuchi T.;
RA "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RA corresponding to the 28.0-40.1 min region on the linkage map.";
RA Mol. Cell. Biol. 16:3363-3377(1996).
RL [3]
RN -!- SIMILARITY: BELONGS TO THE ASL/ATSB FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000247; AAC74570.1; ALT_INIT.
CC DR EMBL; D90791; BAA15168.1; -.
CC DR EMBL; D90792; BAA15171.1; -.
CC DR EcoGene; EG13795; ydem.
CC DR InterPro; IPR007197; Radical_SAM.
CC DR Pfam; PF04055; Radical_SAM; 1.

KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 385 AA; 44518 MW; 964E34F73E680329 CRC64;

Query Match 6.3%; Score 80.5; DB 1; Length 385;
Best Local Similarity 19.4%; Pred. No. 7.4;
Matches 54; Conservative 46; Mismatches 103; Indels 75; Gaps 12;

Qy 10 KVTMSGTVNGHYFEVGEKGRKPK-----YEGQTVRLAV-TKGGGLPFP 51
Db 44 KQYIAASGNQYVFTWQG---GEPTLAGLDFFRKVIHYQORVAGKRIFNALQTNGILLNN 100
Qy 52 AWDLSPQCQVGSIPFTKYPEDIPDYVKQSPFGRYTW-----ERIMNPE-DGAVCTVS 103
Db 101 EWCAFLKEHEFLVIGSIDGQELHVDYRRNSNGTFAKVIAAERLKSQVEENTLTVI 160
Qy 104 NDSSIQCNCFIYH-----VKFSG-----LNFPN--GP 129
Db 161 NNVNHYPLEVHYHFLKSGIKHMQFIELTGTNTIDFSGHSENTFRIDFSVPTAYGK 220
Qy 130 VMQKKTQGWEPN-TERLFARDGMLIGNFMALKLEGGGHYLCEPKSTYKAKKPKVMPG-- 186
Db 221 FMSTIFMQWKNDVGEIFIRQFBSFVSRFL-----GNHTSCIFQESCKDNLVVSNGDI 275
Qy 187 ---YHYVDRLDVTNNHK-DYTSVEQCEISISIAKRPVVA 220
Db 276 YECDFHYVPQYKIGNINKSELKTMNSVOLTAQKKRIPA 313

RESULT 8
RELN_MOUSE
ID_RELN_MOUSE STANDARD; PRT; 3461 AA.
AC Q60841; Q9CUA6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Reelin precursor (EC 3.4.21.-) (Reeler protein).
GN RELN OR Rl.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX TISSUE=Cerebellum;
RX MEDLINE=95231649; PubMed=7715726;
RA D'Arcangelo G., Miao G.G., Chen S.-C., Soares H.D., Morgan J.I.,
RA Curran T.;
RA "A protein related to extracellular matrix proteins deleted in the
RA mouse mutant reeler.";
RA Nature 374:719-723(1995).
RL [2]
RN SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RX MEDLINE=98086481; PubMed=9417911;
RA Royaux I., Lambert de Rouvroit C., D'Arcangelo G., Demirov D.,
RA Goffinet A.M.;
RA "Genomic organization of the mouse reelin gene.";
RA Genomics 46:240-250(1997).
RL [3]
RN SEQUENCE OF 2152-3461 FROM N.A. (ISOFORM 1).
RX STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=95375789; PubMed=7647795;
RA Hirotsune S., Takahara T., Sasaki N., Hirose K., Yoshiki A.,
RA Ohashi T., Kusakabe M., Murakami Y., Muramatsu M., Watanabe S.,
RA Nakao K., Katsuki M., Hayashizaki Y.;
RA "The reeler gene encodes a protein with an EGF-like motif expressed by
RA pioneer neurons.";
RA Nat. Genet. 10:77-83(1995).
RL [4]
RN SEQUENCE OF 3044-3461 FROM N.A. (ISOFORM 2).
RX STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

```

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Guscinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=97141547; PubMed=8987733;
RA D'Arcangelo G., Nakajima K., Miyata T., Ogawa M., Mikoshiba K.,
RA Curran T.;
RT "Reelin is a secreted glycoprotein recognized by the CR-50 monoclonal
RT antibody.";
RL J. Neurosci. 17:23-31(1997).
RN [6]
RP CHARACTERIZATION.
RX MEDLINE=21634304; PubMed=11689558;
RA Quattrocchi C.C., Wamnes F., Persico A.M., Ciafre S.A.,
RA D'Arcangelo G., Farace M.G., Keller F.;
RT "Reelin is a serine protease of the extracellular matrix.";
RL J. Biol. Chem. 277:303-309(2002).
RN [7]
RP TISSUE SPECIFICITY
RX MEDLINE=97325946; PubMed=9182958;
RA Schiffmann S.N., Bernier B., Goffinet A.M.;
RT "Reelin mRNA expression during mouse brain development.";
RL Eur. J. Neurosci. 9:1055-1071(1997).
RN [8]
RP ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RX MEDLINE=99263436; PubMed=10328932;
RA Lambert de Rouvroit C., Bernier B., Royaux I., de Bergueyck V.,
RA Goffinet A.M.;
RT "Evolutionarily conserved, alternative splicing of reelin during brain
RT development.";
RL Exp. Neurol. 156:229-238(1999).
RN [9]
RP BINDING TO VLDLR AND APOER2.
RX MEDLINE=20036019; PubMed=10571241;
RA Hiesberger T., Trommsdorff M., Howell B.W., Goffinet A.M., Mumby M.C.,
RA Cooper J.A., Herz J.;
RT "Direct binding of Reelin to VLDLR receptor and ApoE receptor 2 induces
RT tyrosine phosphorylation of disabled-1 and modulates tau
RT phosphorylation.";
RL Neuron 24:481-489(1999).
RN [10]
RP FUNCTION.
RX MEDLINE=20359755; PubMed=10880573;
RA Yip J.W., Yip Y.P.L., Nakajima K., Capriotti C.;
RT "Reelin controls position of autonomic neurons in the spinal cord.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:8612-8616(2000).
CC -!- FUNCTION: Extracellular matrix serine protease that plays a role
CC in layering of neurons in the cerebral cortex and cerebellum.
CC Regulates microtubule function in neurons and neuronal migration.
CC Affects migration of sympathetic preganglionic neurons in the
CC spinal cord, where it seems to act as a barrier to neuronal
CC migration. Enzymatic activity is important for the modulation of
CC cell adhesion. Binding to the extracellular domains of lipoprotein
CC receptors VLDLR and ApoER2 induces tyrosine phosphorylation of
CC Dab1 and modulation of tau phosphorylation.
CC -!- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2.
CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q060841-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q060841-2; Sequence=VSP_005577;
CC Name=3;
CC IsoId=Q060841-3; Sequence=VSP_005578;
CC -!- TISSUE SPECIFICITY: The major isoform 1 is neuron-specific. It is
CC abundantly produced during brain ontogenesis by the Cajal-Retzius
CC cells and other pioneer neurons located in the telencephalic
CC marginal zone and by granule cells of the external granular layer
CC of the cerebellum. Expression is located in deeper layers in the
CC developing hippocampus and olfactory bulb, low levels of
CC expression are also detected in the immature striatum. At early
CC developmental stages, expressed also in hypothalamic
CC differentiation fields, tectum and spinal cord. A moderate to low
CC level of expression occurs in the septal area, striatal fields,
CC habenular nuclei, some thalamic nuclei, particularly the lateral
CC geniculate, the retina and some nuclei of the reticular formation
CC in the central field of the medulla. Very low levels found in
CC liver and kidney. No expression in radial glial cells, cortical
CC plate, Purkinje cells and inferior olivary neurons. The minor
CC isoform 2 is only expressed in non neuronal cells. The minor
CC isoform 3 is found in the same cells as isoform 1, but is almost
CC undetectable in retina and brain stem.
CC -!- DEVELOPMENTAL STAGE: First detected at embryonic day 11.5.
CC Expression increases up to birth and remains high from post-natal
CC day 2 to 11 in both cerebellum and fore/midbrain. Expression
CC declines thereafter and is largely brain specific in the adult.
CC -!- DOMAIN: The basic C-terminal region is essential for secretion.
CC -!- PTM: N-glycosylated and to a lesser extent also O-glycosylated.
CC -!- DISEASE: Defects in reelin are the cause of the autosomal recessive
CC reeler (rl) phenotype which is characterized by impaired motor
CC coordination, tremors and ataxia. Neurons in affected mice fail to
CC reach their correct locations in the developing brain, disrupting
CC the organization of the cerebellar and cerebral cortices and other
CC laminated regions.
CC -!- SIMILARITY: Belongs to the reelin family.
CC -!- SIMILARITY: Contains 8 EGF-like domains.
CC -!- SIMILARITY: Contains 15 BNR repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U24703; AAB91599.1; -;
CC EMBL; D63520; BAA09788.1; ALT_INIT.
CC EMBL; AK017094; BAB30592.1; -;
CC MGD; MGI:103022; ReIn.
CC GO; GO:0005615; C:extracellular space; IDA.
CC GO; GO:0007420; P:brain development; IMP.
CC GO; GO:0016477; P:cell migration; IMP.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR002860; GH_BNR.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR002861; Reeler.
CC Pfam; PF02012; BNR; 15.
CC Pfam; PF00008; EGF; 3.
CC Pfam; PF02014; Reeler; 1.
CC SMART; SM00181; EGF; 5.
CC PROSITE; PS00022; EGF_1; 7.
CC PROSITE; PS01186; EGF_2; 6.
CC PROSITE; PS00026; EGF_3; 5.
CC Hydrolase; Serine protease; Developmental protein; Matrix protein;
CC Cell adhesion; EGF-like domain; Glycoprotein; Repeat; Signal;
CC Alternative splicing.
CC SIGNAL 1 26
CC CHAIN 27 3461
CC FT
CC POTENTIAL.
CC REELIN.

```
FT DOMAIN 40 172 REELER.
FT DOMAIN 671 702 EGF-LIKE 1.
FT DOMAIN 1030 1061 EGF-LIKE 2.
FT DOMAIN 1409 1442 EGF-LIKE 3.
FT DOMAIN 1765 1796 EGF-LIKE 4.
FT DOMAIN 2129 2161 EGF-LIKE 5.
FT DOMAIN 2478 2509 EGF-LIKE 6.
FT DOMAIN 2853 2884 EGF-LIKE 7.
FT DOMAIN 3228 3260 EGF-LIKE 8.
FT REPEAT 593 604 BNR 1.
FT REPEAT 799 810 BNR 2.

Query Match 6.3%; Score 80.5; DB 1; Length 3461;
Best Local Similarity 19.9%; Pred. No. 99;
Matches 56; Conservative 25; Mismatches 75; Indels 125; Gaps 12;

Qy 12 YMSGTVNGHYPEVGGDK-----GKPYEGEQTVRLAVTKG-----GPLPPAWDI-----LSPQ 59
Db 2080 YYAGTTQGRREVVFHGLHLCG-----SVRFWYQGFYPAGSQPVTAIDNVYIGQP 2132

Qy 60 CQ---YG-----SIPFTKYPEDIPDYVKQSPGGRYTWERIMNF----- 94
Db 2133 CEEMCYHGSGINGTKICIDPGYSGPTCKISTKNPDKDFEGQLESDRFLMGGGKPS 2192

Qy 95 -----EDGAVCTVNDSSIQNCFTYHYVKF-----SGLNFPNPGP 129
Db 2193 RKCGLSSGNLFFNEDGLRLVTRDLS-----HARFVQFFMRLCGKGVDPDRSQP 2246

Qy 130 VM-----VLTQVSLNGSLWSLLOFLFNSNVGRYIALEMPKARSGSTRLRWQPSENGHFYSP 2306
Db 2247 VLLQVSLNGSLWSLLOFLFNSNVGRYIALEMPKARSGSTRLRWQPSENGHFYSP 2306

Qy 147 -ARQMLIGNFMALKBGGGHYLCFEPKSTYKAKKPVKMPG 186
Db 2307 WVIDQILIGNI-----SGNTVLEDDFSTLDSRKWLHPG 2341
```

RESULT 9

```
SYA_STAAM
ID SYA_STAAM STANDARD; PRT; 876 AA.
AC Q99TNI;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (AlaRS).
CN ALAS OR SAV1618 OR SRA146.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240 (2001).
CC -!- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
CC diphosphate + L-alanyl-tRNA(Ala).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

```
CC -----
CC EMBL; AP003362; BAB57780.1; -
CC DR EMBL; AP003362; BAB57780.1; -
CC DR PIR; A89944; A89944.
CC DR HAMAP; MF 00036; -; 1.
CC DR InterPro; IPR003156; DHHA1.
CC DR InterPro; IPR002318; tRNA-synt_2c.
CC DR InterPro; IPR006193; tRNA-synt_Ala.
CC DR Pfam; PF02272; DHHA1; 1.
CC DR Pfam; PF01411; tRNA-synt_2c; 1.
CC DR PRINTS; PRO0980; TRNASYNTHALA.
CC DR TIGRFAMs; TIGR00344; alas; 1.
CC DR PROSITE; PS50860; AA tRNA_LIGASE II ALA; 1.
CC KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC Complete proteome.
CC SQ SEQUENCE 876 AA; 98538 MW; 2B2BC79041AC264F CRC64;
```

```
Query Match 6.2%; Score 79.5; DB 1; Length 876;
Best Local Similarity 21.4%; Pred. No. 24;
Matches 43; Conservative 25; Mismatches 56; Indels 77; Gaps 10;

Qy 16 TVNGH-YFEVGDGKPKPYEGEQTVRLAVTKGGLPFAWDILSPQCQVGSIP-----FTKY 70
Db 81 TARHHTPEMLGNFSIGDYFKQEAIE-----FAWEFLTSDKWMGMEPDKLYVTIH 130

Qy 71 PEDIPDYVKQSPGGRYTWERIMNFEDGAVCTVNDSSIQNCFTYHYVKFSGLNFPNPGPV 130
Db 131 PEDMEAY-----NIWKDIDGLEESRII-----RIEGLN-----FWDIGEGPSG-- 167

Qy 131 MOKKTOGWEPNTERLFAKDGMIGNFMALKBGGGHYLCFEPKSTYKAKKPVKMPGYHYV 190
Db 168 -----PNTETFYDR-GEAYGQDDPAEEMYPGEN----- 195

Qy 191 DRKLDV-----TNHKNKDYT 204
Db 196 ERYLEVWNLVSEFNNHNDHS 216
```

RESULT 10

```
SYA_STAAM
ID SYA_STAAM STANDARD; PRT; 876 AA.
AC Q8NW87;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (AlaRS).
CN ALAS OR MW1568.
OS Staphylococcus aureus (strain MW2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA."
RT Lancet 359:1819-1827 (2002).
CC -!- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
CC diphosphate + L-alanyl-tRNA(Ala).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP004827; BAB95433.1; -.
DR HAMAP; MF_00036; -.
DR InterPro; IPR003156; DHHA1.
DR InterPro; IPR002318; trna-synt_2c.
DR InterPro; IPR006193; trna-synt_ala.
DR Pfam; PF022272; DHHA1; 1.
DR Pfam; PF01411; trna-synt_2c; 1.
DR PRINITS; PR00980; TRNASYNTHALA.
DR TIGRFAMs; TIGR00344; alas; 1.
DR PROSITE; PS50860; AA TRNA LIGASE II ALA; 1.
KW Aminoacyl-tRNA synthetase; Protein Biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
SQ SEQUENCE 876 AA; 98505 MW; 5D9D662D8DADDFC CRC64;

Query Match 6.2%; Score 79.5; DB 1; Length 876;
Best Local Similarity 21.4%; Pred. No. 24;
Matches 43; Conservative 25; Mismatches 56; Indels 77; Gaps 10;

QY 16 TVNGH-YEVEGDGKGYEGEQTVRLAVTKGGLPFPFADWILSPQCYGSIP-----FTKY 70
DB 81 TARHHTFFEMGLNFGISGYFFKQEAIE-----FAWEFLTSDKXWGMEDPKLYVTIH 130
QY 71 PEDIPDYVKQSPPGRYTWERIMNFEDGAVCTVSDSSIQGNCFIYHVKFSGLNFPNGPV 130
DB 131 PEDMEAY-----NIHKKDGLIESRII-----RIEGN-----FWDIGEGPSG-- 167
QY 131 MOKTKQGWEPTNTERLPARDGMLIGNFNWALKEGGHYLCPEKSYTKAKKPKVMPGYHYV 190
DB 168 -----PNTIEFYDR-GEAYGQDDPAEEMYPGGEN----- 195
QY 191 DRKLDV-----TNHNDKYT 204
DB 196 BRYLEVWNLVTFSEFNHNDHS 216

RESULT 11
ID H136 ARATH STANDARD; PRT; 403 AA.
AC O82660;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Photosystem II stability/assembly factor HCF136, chloroplast
DE Precursor.
GN HCF136 OR AT5G23120 OR MJ24.11.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Wassilewskija; TISSUE=Leaf;
RX MEDLINE=98409534; PubMed=9736608;
RA Meurer J., Plucken H., Kowallik K.V., Westhoff P.;
RT "A nuclear-encoded protein of prokaryotic origin is essential for the
RT stability of photosystem II in Arabidopsis thaliana.";
RL EMBO J. 17:5286-5297(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20181125; PubMed=10718197;
RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asanizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.
RT Sequence features of the regions of 1,044,062 bp covered by thirteen
RT physically assigned P1 clones.";
RN DNA Res. 4:291-300(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;

```

```

RX MEDLINE=22954850; PubMed=14593172;
RA Yamada K., Lim J., Dale J.M., Chen H., Shinn P., Palm C.J.,
RA Southwick A.M., Wu H.C., Kim C.J., Nguyen M., Pham P.K., Cheuk R.F.,
RA Karlin-Newmann G., Liu S.X., Lam B., Sakano H., Wu T., Yu G.,
RA Miranda M., Quach H.L., Tripp M., Chang C.H., Lee J.M., Toriumi M.J.,
RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA Arakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA Khan S., Koesema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Satou M., Tamse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Empirical analysis of transcriptional activity in the Arabidopsis
RT genome.";
RL Science 302:842-846(2003).
CC -!- FUNCTION: Essential for photosystem II (PSII) biogenesis; putative
CC protein stability or assembly factor for PSII.
CC -!- SUBCELLULAR LOCATION: Chloroplast; within the thylakoid lumen but
CC attached to the membrane. Restricted to the stromal lamellae.
CC -!- TISSUE SPECIFICITY: Expression in green tissue, not roots.
CC -!- DEVELOPMENTAL STAGE: Accumulates also in dark-grown seedlings.
CC -!- SIMILARITY: Belongs to the ycf48 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y15628; CAA75723.1; -.
CC EMBL; AB006708; BAB09829.1; -.
CC EMBL; AY045691; AAK74049.1; -.
CC PIR; T51828; T51828.
CC InterPro; IPR002860; GH_BNR.
CC InterPro; IPR006311; TaT.
CC Pfam; PF02012; BNR; 4.
CC TIGRFAMs; TIGR01409; TAT_signal_seq; 1.
KW Chloroplast; Thylakoid; Transit peptide; Photosystem II; Membrane.
FT TRANSIT 1 53 CHLOROPLAST (POTENTIAL).
FT TRANSIT 54 78 THYLAKOID (POTENTIAL).
FT CHAIN 79 403 PHOTOSYSTEM II STABILITY/ASSEMBLY FACTOR
FT HCF136.
SQ SEQUENCE 403 AA; 44103 MW; 11079552F917FF9D CRC64;

Query Match 6.1%; Score 79; DB 1; Length 403;
Best Local Similarity 23.9%; Pred. No. 11;
Matches 49; Conservative 25; Mismatches 89; Indels 42; Gaps 11;

QY 18 NGHYFEVGDGKGYEGEQTVRLAVTKGGLPFPFADWILSPQCY-GSIPFTKYPED---- 73
DB 144 NYRFNSISFGKGEWIIKGPAILLYTADAGE---NWDRIPLSSQLPGDMVFYKATEDKSA 200
QY 74 --IPD-----YVKQSPFGRYTWERIMNFEDGAVCTVSDSSIQGNCFIYHVKFSGLNFPN 127
DB 201 EMTVDEGAIYVTSN--RGYNKAAIQETVSATLNRTVSSGISGASY-YGTFTSAVNRSPD 257
QY 128 GPVMQKKTQG-----WEPNT-----ERLFARDGMLIG-----NNFMALKLEGGGHYIC--- 170
DB 258 GRYVAVSSRGNFLLTWEPGQPVQWPHNRAVARRIQNMGRADGGLWLLVRGGGLYLSKGT 317
QY 171 ----EFKSTYAKKPKVMPGVHYVD 191
DB 318 GITEEFEEV-----PVQSRGFGILD 337

RESULT 12
CYBC_BRAJA
ID CYBC_BRAJA STANDARD; PRT; 687 AA.
AC P51131;
DT 01-OCT-1996 (Rel. 34, Created)

```


RP SEQUENCE OF 338-898 FROM N.A.
 RX MEDLINE=88166734; PubMed=3350013;
 RA Laum N., Wang Y., Mathews C.K., Rueger W.;
 RT "Deoxycytidylate hydroxymethylase gene of bacteriophage T4.
 RL Nucleotide sequence determination and over-expression of the gene.";
 RN Eur. J. Biochem. 172:553-563(1988).
 [5]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 1-388.
 RX MEDLINE=96292335; PubMed=8679362;
 RA Wang J., Yu P., Lin T.C., Konigsberg W.H., Steitz T.A.;
 RT "Crystal structures of an NH2-terminal fragment of T4 DNA polymerase
 and its complexes with single-stranded DNA and with divalent metal
 ions.";
 RL Biochemistry 35:9110-9119(1996).
 CC -!- FUNCTION: This polymerase possesses two enzymatic activities: DNA
 synthesis (polymerase) and an exonucleolytic activity that
 degrades single stranded DNA in the 3' to 5' direction.
 CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 + {DNA} (N).
 CC -!- SIMILARITY: Belongs to the DNA polymerase type-B family.
 CC -!- DATABASE: NAME=Worthington enzyme manual;
 WWW="http://www.worthington-biochem.com/DNAPT4/".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M10160; AAC05397.1; -;
 DR EMBL; X00769; CAA25344.1; -;
 DR EMBL; AF158101; AAD42468.1; -;
 DR EMBL; M37159; AAA21706.1; -;
 DR PIR; J50791; DJBPT4.
 DR PDB; 1NOY; 14-OCT-96.
 DR PDB; 1NCZ; 14-OCT-96.
 DR InterPro; IPR006172; DNA pol B.
 DR InterPro; IPR006134; DNA pol B dom.
 DR InterPro; IPR006133; DNA pol B_exo.
 DR Pfam; PF00136; DNA pol B; 1.
 DR Pfam; PF03104; DNA pol_B_exo; 1.
 DR PRINTS; PR00106; DNAPOLB.
 DR SMART; SM00486; POLBc; 1.
 DR PROSITE; P500116; DNA POLYMERASE B; 1.
 KW Transferrase; DNA-directed DNA polymerase; DNA replication; Hydrolase;
 KW Exonuclease; DNA-binding; 3D-structure.
 FT CONFLICT 89 89 A -> V (IN REF. 2; CAA25344).
 FT STRAND 4 11
 FT TURN 12 13
 FT STRAND 14 21
 FT TURN 22 24
 FT STRAND 25 31
 FT STRAND 37 38
 FT HELIX 40 42
 FT STRAND 58 59
 FT HELIX 63 67
 FT TURN 68 71
 FT TURN 83 83
 FT TURN 84 94
 FT HELIX 103 105
 FT STRAND 108 114
 FT TURN 123 125
 FT STRAND 132 137
 FT TURN 138 141
 FT STRAND 142 148
 FT TURN 149 150
 FT TURN 152 153
 FT TURN 161 165
 FT TURN 168 171
 FT HELIX 177 181
 FT TURN 182 182

FT STRAND 183 188
 FT HELIX 191 204
 FT STRAND 209 211
 FT TURN 215 218
 FT HELIX 219 229
 FT TURN 230 231
 FT TURN 233 235
 FT HELIX 236 238
 FT TURN 241 242
 FT STRAND 245 248
 FT STRAND 259 262
 FT TURN 263 264
 FT STRAND 266 267
 FT HELIX 270 277
 FT HELIX 287 295
 FT HELIX 307 313
 FT TURN 314 315
 FT HELIX 316 335
 FT HELIX 337 348
 FT TURN 349 349
 FT TURN 352 356
 FT HELIX 358 368
 FT TURN 369 369
 SQ SEQUENCE 898 AA; 103609 MW; 925300C4CA5C7A24 CRC64;
 Query Match 6.1%; Score 79; DB 1; Length 898;
 Best Local Similarity 22.3%; Pred. No. 27;
 Matches 49; Conservative 37; Mismatches 66; Indels 68; Gaps 14;
 QY 41 LAVTKGGLPFADILSPQCQVGSIPFTKYPED---IP---DYVKQSPFGRYTWE----- 89
 Db 342 LMSYVAKMPFS-GVMSPIKTWDALIFNSLKGHEKVIQQGSHVQKSPFGAFVFPKPIA 400
 QY 90 --RINNF-----DGAVCTVS-NDSSIQGNCHYI--HVKFSG-----LNFPNGPVM 131
 Db 401 RRYIMSFDTLSLYPSIIRQVNISPTIRGQFKVHPHEIYAGTAPKPSDEYSCSPNGWY 460
 QY 132 OKKTQGWEPN-----TERLFARDGMLGNFMALK---LEGSGHYLCFFKSY 176
 Db 461 DKHQGIIPKEIAKVFFORKWKKWFABE-----MNAEAIKKIIMKGAGS--CSTKPEV 513
 QY 177 KAKFVKMPGXYHVDKLDVTNNKDYTS-----VEQCE 210
 Db 514 E-----RYVKFSDDFLNELSNYTESVLSLIECE 543
 RESULT 14
 ID LOXC ARATH STANDARD; PRT; 896 AA.
 AC P38418; Q8W4E4; Q9MIU5;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Lipoxigenase, chloroplast precursor (EC 1.13.11.12).
 GN LOX2 OR AT3G45140 OR T14D3.80.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eumecids II; Brassicales; Brassicaceae; Arabidopsi.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94120003; PubMed=8290626;
 RA Bell E., Mullet J.E.;
 RT "Characterization of an Arabidopsis lipoxigenase gene responsive to
 methyl jasmonate and wounding.";
 RL Plant Physiol. 103:1133-1137(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016720; PubMed=11130713;
 RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unsel M.,
 RA Fartmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,

[illegible]

Db 2248 VILQYSLNGLSWSLLQEFLEFSNSSNVGRYIALEMPLEKARSGSTLRWWQPSENGHFYSP 2307
QY 147 -ARDGMLIGNNFMAKLEGGGHYLCEPKSTYKAKKPVKMPG 186
Db 2308 WVIDQILIGGNI-----SGNTVLEDDFSTLDSRAWLLHPG 2342

Search completed: August 12, 2004, 06:20:09
Job time : 15.4467 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 06:12:47 ; Search time 61.1578 Seconds
(without alignments)
1212.385 Million cell updates/sec

Title: US-09-890-463-4
Perfect score: 1287
Sequence: 1 SVIAQMTYKYVMGTVNGH.....KPVVACRFVRVKSRRKYAVA 235

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp archaea:*
2: sp bacteria:*
3: sp fungi:*
4: sp human:*
5: sp invertebrate:*
6: sp mammal:*
7: sp mhc:*
8: sp organelle:*
9: sp phage:*
10: sp plant:*
11: sp rodent:*
12: sp virus:*
13: sp vertebrate:*
14: sp unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1156	89.8	221 5 Q95P04	Q95P04 gonipora t
2	783	60.8	225 5 Q9U6Y8	Q9U6Y8 discosoma s
3	750	58.3	230 5 Q9GTJ7	Q9GTJ7 discosoma s
4	737.5	57.3	232 5 Q9U6Y7	Q9U6Y7 discosoma s
5	707	54.9	236 5 Q8T6U0	Q8T6U0 dendroneph
6	706	54.9	225 5 Q9G3P5	Q9G3P5 montastraea
7	701	54.5	225 5 Q7Z0W4	Q7Z0W4 montastraea
8	674	52.4	225 5 Q7Z0W5	Q7Z0W5 montastraea
9	671	52.1	266 5 Q9U6Y3	Q9U6Y3 clavularia
10	669.5	52.0	227 5 Q7Z0W6	Q7Z0W6 montastraea
11	667.5	51.9	225 5 Q7Z0W9	Q7Z0W9 montastraea
12	667.5	51.9	227 5 Q9G2P9	Q9G2P9 montastraea
13	667.5	51.9	227 5 Q7Z0W8	Q7Z0W8 montastraea
14	666.5	51.8	225 5 Q95UA7	Q95UA7 montastraea
15	656.5	51.0	225 5 Q8T5F1	Q8T5F1 montastraea
16	654.5	50.9	227 5 Q95VT0	Q95VT0 montastraea

17	654.5	50.9	234	5	Q8T5F2	Q8T5f2 montastraea
18	654	50.8	224	5	Q8MU48	Q8mu48 montastraea
19	633.5	50.8	234	5	Q7Z0W7	Q7z0w7 montastraea
20	631.5	49.1	225	5	Q816J8	Q816j8 trachyphyl
21	625.5	48.6	234	5	Q8MU47	Q8mu47 montastraea
22	612.5	47.6	259	5	Q8MMA2	Q8mma2 agaricia fr
23	609.5	47.4	231	5	Q8T5E9	Q8t5e9 ricordea fl
24	601	46.7	231	5	Q8T5E8	Q8t5e8 ricordea fl
25	600.5	46.7	231	5	Q8T6T8	Q8t6t8 discosoma s
26	586.5	45.6	231	5	Q81SF8	Q81sf8 parasicyoni
27	579.5	45.0	239	5	Q8MMAL	Q8mmal agaricia ag
28	578.5	44.9	231	5	Q8MU46	Q8mu46 ricordea fl
29	574.5	44.6	232	5	Q9GP15	Q9gp15 anemonia su
30	569.5	44.3	227	5	Q95W86	Q95w86 condylactis
31	565.5	43.9	227	5	Q95W85	Q95w85 radianthus
32	561.5	43.6	227	5	Q95W11	Q95w11 condylactis
33	561.5	43.6	232	5	Q9GZ28	Q9gz28 anemonia su
34	561	43.6	228	5	Q9GP16	Q9gp16 anemonia su
35	549.5	42.7	227	5	Q8MU45	Q8mu45 condylactis
36	536.5	41.7	235	5	Q8T5P0	Q8t5f0 scolymia cu
37	534	41.5	214	5	Q86LV7	Q86lv7 meandrina m
38	534	41.5	228	5	Q86LV4	Q86lv4 radianthus
39	531	41.3	229	5	Q9U6Y6	Q9u6y6 anemonia ma
40	529.5	41.1	225	5	Q8T6T9	Q8t6t9 radianthus
41	529	41.1	214	5	Q86LV8	Q86lv8 meandrina m
42	527.5	41.0	234	5	Q8T5E3	Q8t5f3 scolymia cu
43	516.5	40.1	229	5	Q8T5E7	Q8t5e7 condylactis
44	488.5	38.0	231	5	Q9U6Y5	Q9u6y5 zoanthus sp
45	480.5	37.3	231	5	Q9U6Y4	Q9u6y4 zoanthus sp

ALIGNMENTS

RESULT 1

Q95P04 PRELIMINARY; PRT; 221 AA.
AC Q95P04;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE GFP-like chromoprotein.
OS Gonipora tenuidens.
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Fungiina; Poritidae; Gonipora.
OX NCBI_TaxID=75301;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21538626; PubMed=11682051;
RA Gurskaya N.G., Fradkov A.F., Tersikh A., Matz M.V., Labas Y.A.,
RA Martynov V.I., Yanushevich Y.G., Lukyanov K.A., Lukyanov S.A.;
RT "GFP-like chromoproteins as a source of far-red fluorescent
RT proteins(1).";
RL FEBS Lett. 507:16-20(2001).
DR EMBL; AF383156; AAL27542.1; --
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR ProDom; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 221 AA; 24918 NW; 93P9F4B5C2003CB4 CRC64;

Query Match 89.8%; Score 1156; DB 5; Length 221;
Best Local Similarity 96.4%; Pred. No. 2.1e-100; Indels 0; Gaps 0;
Matches 212; Conservative 3; Mismatches 5;
QY 1 SVIAQMTYKYVMGTVNGHYFEVGDGKGYEGEQTVRLAVTKGGLPFPFADILSPQC 60
Db 2 SVIAQMTYKYVMGTVNGHYFEVGDGKGYEGEQTVRLAVTKGGLPFPFADILSPQS 61
QY 61 QYGSIPFTKYPEDIPDYVKQSPFGYTWERINMFEDGAVCTVNSDSSIQGNCFIYHVKFS 120
Db 62 QYGSIPFTKYPEDIPDYVKQSPFGYTWERINMFEDGAVCTVNSDSSIQGNCFIYHVKFS 121


```
Query Match          57.3%; Score 737.5; DB 5; Length 232;
Best Local Similarity 60.2%; Pred. No. 4.2e-61;
Matches 130; Conservative 38; Mismatches 47; Indels 1; Gaps 1;

QY 1 SVIAKQMTYKVMYSGTVNGHYHFEVGDGKGKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
Db 6 SVIKEMLIDLHLEGTENGHYHFEIKGKGQPNEGTNTVTLVTKGGPLPFGWHILCPQF 65

QY 61 QYGSIPFTKYPEDIDPYKQSPGRTYWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKFS 120
Db 66 QYGNKAFVHHPNIDHLYKLSPGRTYWERSMHFEDGGLCCITNDISLTGNCFFYDIKFT 125

QY 121 GLNFPNGPVMOKKTQGWEPNTERLFARDGMLIGNFMALKEGGHYLCYCEFKSTYKAKK 180
Db 126 GLNFPNGPVMOKKTQGWEPNTERLFARDGMLIGNFMALKEGGHYLCYCEFKSTYKAKK 185

QY 181 -PVKMPGYHYVDKLDVTNNHNDYTSVQCEISAR 215
Db 186 AALKMPGYHYVDKLDVTNNHNDYTSVQCEISAR 221

RESULT 5
Q8T6U0 PRELIMINARY; PRT; 236 AA.
AC Q8T6U0
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Green fluorescent protein.
OS Dendronephthya sp. SSAL-2002.
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Alcyonaria; Alcyonacea;
OC Nephtheidae; Dendronephthya.
OX NCBI_TaxID=191210;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21927629; PubMed=11929996;
RA Labas Y.A., Gurskaya N.G., Yanushkevich Y.G., Fradkov A.F.,
RA Lukyanov K.A., Lukyanov S.A., Matz M.V.;
RT "Diversity and evolution of the green fluorescent protein family.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4256-4261(2002).
DR EMBL; AF420591; AAM10625.1; -.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP like.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFP.
DR ProDom; PD013756; Green fl protein; 1.
SQ SEQUENCE 236 AA; 26840 MW; CE1707CFF9334A90 CRC64;

Query Match          54.9%; Score 707; DB 5; Length 236;
Best Local Similarity 55.8%; Pred. No. 3.1e-58;
Matches 120; Conservative 45; Mismatches 50; Indels 0; Gaps 0;

QY 1 SVIAKQMTYKVMYSGTVNGHYHFEVGDGKGKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
Db 2 NLIKEDMRVKVHMEGNVNGHAFVIEGEGKGRPYEGTQTLNLTVKEGAPLPFSYDILTTAL 61

QY 61 QYGSIPFTKYPEDIDPYKQSPGRTYWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKFS 120
Db 62 HYGNVFTYEPADITDYKQSPGRTYWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKFS 121

QY 121 GLNFPNGPVMOKKTQGWEPNTERLFARDGMLIGNFMALKEGGHYLCYCEFKSTYKAKK 180
Db 122 GNFPPNGPVMOKKTQGWEPNTERLFARDGMLIGNFMALKEGGHYLCYCEFKSTYKAKK 181

QY 181 PVKMPGYHYVDKLDVTNNHNDYTSVQCEISAR 215
Db 182 VVQLPDYHFDHRIELSDNSDYNKVKLYEHGVAR 216

RESULT 6
Q963F5 PRELIMINARY; PRT; 225 AA.
ID Q963F5
```

```
Q963F5;
AC 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Green fluorescent protein.
OS Montastraea cavernosa (great star coral).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Faviina; Faviidae; Montastraea.
OX NCBI_TaxID=63558;
RN [1]
RP SEQUENCE FROM N.A.
RA Lesser M.P., Barry T.M., Matz M.V., Lukyanov S.A.,
RA Falkowski P., Gorbunov M., Kolber Z.;
RT "Green fluorescent proteins in Caribbean Scleractinian corals.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF384683; AAK62982.2; -.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP like.
DR InterPro; IPR00786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFP.
DR ProDom; PD013756; Green fl protein; 1.
SQ SEQUENCE 225 AA; 25847 MW; 77DE7D7C616929AF CRC64;

Query Match          54.9%; Score 706; DB 5; Length 225;
Best Local Similarity 56.7%; Pred. No. 3.6e-58;
Matches 122; Conservative 45; Mismatches 46; Indels 0; Gaps 0;

QY 1 SVIAKQMTYKVMYSGTVNGHYHFEVGDGKGKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
Db 2 SVIKPIMEIKLRMQGVNGHFKPVKGEKGKPFEGTQTLNLTVKEGAPLPFAWDILTSF 61

QY 61 QYGSIPFTKYPEDIDPYKQSPGRTYWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKFS 120
Db 62 QYGNRVFTKYDPDIDPYKQSPGRTYWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKFS 121

QY 121 GLNFPNGPVMOKKTQGWEPNTERLFARDGMLIGNFMALKEGGHYLCYCEFKSTYKAKK 180
Db 122 GNFPPNGPVMOKKTQGWEPNTERLFARDGMLIGNFMALKEGGHYLCYCEFKSTYKAKK 181

QY 181 PVKMPGYHYVDKLDVTNNHNDYTSVQCEISAR 215
Db 182 RVQLPDYHFDHRIELSDNSDYNKVKLYEHGVAR 216

RESULT 7
Q7ZOW4 PRELIMINARY; PRT; 225 AA.
AC Q7ZOW4
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Green fluorescent protein.
OS Montastraea cavernosa (great star coral).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Faviina; Faviidae; Montastraea.
OX NCBI_TaxID=63558;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=mc6;
RX MEDLINE=22689801; PubMed=12777529;
RA Kelman I.V., Matz M.V.;
RT "Molecular Basis and Evolutionary Origins of Color Diversity in Great
Star Coral Montastraea cavernosa (Scleractinia: Faviida).";
RL Mol. Biol. Evol. 20:1125-1133(2003).
DR EMBL; AY181557; AAO61603.1; -.
SQ SEQUENCE 225 AA; 25827 MW; A600ADD716C5921E CRC64;

Query Match          54.5%; Score 701; DB 5; Length 225;
Best Local Similarity 56.3%; Pred. No. 1.1e-57;
Matches 121; Conservative 42; Mismatches 52; Indels 0; Gaps 0;

QY 1 SVIAKQMTYKVMYSGTVNGHYHFEVGDGKGKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
```

```
Db 2 SVIKPDMKIKLRMEGAVNGHNFVIEGEGKGPFEQTQINLTIVKGGPLPAYDILTTAF 61
Qy 61 QYGSIPFTKYPEDIPDYVKQSPGPGRYTWERIMNFDGAVCTVSDSSIOGNCFIYHVKFS 120
Db 62 QYGNRAFTKYPEDIPDYVKQSPGPGRYTWERIMNFDGAVCTVSDSSIOGNCFIYHVKFS 121
Qy 121 GLNFPNGPVQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCBFKSTYKAKK 180
Db 122 GVNFPNGPVQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCBFKSTYKAKK 181
Qy 181 PVKMPGHHYVDRKLDVTHNNDKDYTSVEQCEISIAI 215
Db 182 RVQLPDYHFVDHRIELSHDNDYNTVKLSNAEAR 216
```

```
RESULT 8
Q7Z0W5 PRELIMINARY; PRT; 225 AA.
ID Q7Z0W5
AC Q7Z0W5;
DT 01-OCT-2003 (TremBLrel. 25, Created)
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Cyan fluorescent protein.
OS Montastraea cavernosa (great star coral).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Faviida; Faviidae; Montastraea.
OX NCBI_TaxID=63558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mc5;
RX MEDLINE=22689801; PubMed=12777529;
RA Kelmanson I.V., Matz M.V.;
RT "Molecular Basis and Evolutionary Origins of Color Diversity in Great Star Coral Montastraea cavernosa (Scleractinia: Faviida).";
RL Mol. Biol. Evol. 20:1125-1133(2003).
DR EMBL; AY181556; AAO61602.1; -.
SQ SEQUENCE 225 AA; 25843 MW; 13708587B7D93E35 CRC64;
```

```
Query Match 52.4%; Score 674; DB 5; Length 225;
Best Local Similarity 54.9%; Pred. No. 3.7e-55;
Matches 118; Conservative 42; Mismatches 55; Indels 0; Gaps 0;

Qy 1 SVIAKQMTYKVMGTVNGHYFEVGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
Db 2 SVIKSVMIKILHMDGIVNGHKFMITGEGKGPFEQTHIILKVBEGGPLPAYDILTTAF 61
Qy 61 QYGSIPFTKYPEDIPDYVKQSPGPGRYTWERIMNFDGAVCTVSDSSIOGNCFIYHVKFS 120
Db 62 QYGNRVFTKYPEDIPDYVKQSPGPGRYTWERIMNFDGAVCTVSDSSIOGNCFIYHVKFS 121
Qy 121 GLNFPNGPVQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCBFKSTYKAKK 180
Db 122 GVNFPNGPVQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCBFKSTYKAKK 181
Qy 181 PVKMPGHHYVDRKLDVTHNNDKDYTSVEQCEISIAI 215
Db 182 GWLPEYHFVDHRIELSHDNDYNTVEYVENAVAR 216
```

```
RESULT 9
Q9U6Y3 PRELIMINARY; PRT; 266 AA.
ID Q9U6Y3
AC Q9U6Y3;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Fluorescent protein FP484.
OS Clavularia sp.
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Alcyonaria; Alcyonacea;
OC Clavulariidae; Clavularia.
OX NCBI_TaxID=86521;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=99436614; PubMed=10504696;
RA Matz M.V., Fradkov A.F., Labas Y.A., Savitsky A.P., Zaraisky A.G.,
RT Markelov M.L., Lukyanov S.A.;
RT "Fluorescent proteins from nonbioluminescent Anthozoa species.";
RL Nat. Biotechnol. 17:969-973(1999).
DR EMBL; AF168424; AAF03374.1; -.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFP; 1.
DR ProDom; PD013756; Green fl protein; 1.
SQ SEQUENCE 266 AA; 30450 MW; B4E97406E2708854 CRC64;
```

```
Query Match 52.1%; Score 671; DB 5; Length 266;
Best Local Similarity 55.6%; Pred. No. 8.6e-55;
Matches 119; Conservative 37; Mismatches 55; Indels 0; Gaps 0;

Qy 2 VIAKQMTYKVMGTVNGHYFEVGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 61
Db 45 VIKPDMKIKLRMEGAVNGHNFVIEGEGKGPFEQTHIILKVBEGGPLPAYDILTTAF 104
Qy 62 YGSIPFTKYPEDIPDYVKQSPGPGRYTWERIMNFDGAVCTVSDSSIOGNCFIYHVKFS 121
Db 105 YGNRALTKYPDDIADYFKQSPGPGRYTWERIMNFDGAVCTVSDSSIOGNCFIYHVKFS 164
Qy 122 LNFPNGPVQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCBFKSTYKAKK 181
Db 165 MNFPNGPVQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCBFKSTYKAKK 224
Qy 182 VMFPGHYVDRKLDVTHNNDKDYTSVEQCEISIAI 215
Db 225 VKLPDYHFVDHRIELSHDNDYNTVEYVENAVAR 258
```

```
RESULT 10
Q7Z0W6 PRELIMINARY; PRT; 227 AA.
ID Q7Z0W6
AC Q7Z0W6;
DT 01-OCT-2003 (TremBLrel. 25, Created)
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Green fluorescent protein.
OS Montastraea cavernosa (great star coral).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Faviida; Faviidae; Montastraea.
OX NCBI_TaxID=63558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mc4;
RX MEDLINE=22689801; PubMed=12777529;
RA Kelmanson I.V., Matz M.V.;
RT "Molecular Basis and Evolutionary Origins of Color Diversity in Great Star Coral Montastraea cavernosa (Scleractinia: Faviida).";
RL Mol. Biol. Evol. 20:1125-1133(2003).
DR EMBL; AY181555; AAO61601.1; -.
SQ SEQUENCE 227 AA; 26055 MW; 4BE2CB64FDB0B890 CRC64;
```

```
Query Match 52.0%; Score 669.5; DB 5; Length 227;
Best Local Similarity 54.9%; Pred. No. 9.8e-55;
Matches 117; Conservative 41; Mismatches 52; Indels 3; Gaps 1;

Qy 1 SVIAKQMTYKVMGTVNGHYFEVGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
Db 2 SVIKPDMKIKLRMEGAVNGHNFVIEGEGKGPFEQTHIILKVBEGGPLPAYDILTTAF 61
Qy 61 QYGSIPFTKYPEDIPDYVKQSPGPGRYTWERIMNFDGAVCTVSDSSIOGNCFIYHV 117
Db 62 DYGNRVFAKYPEDIPDYFKQSPGPGRYTWERIMNFDGAVCTVSDSSIOGNCFIYHV 121
Qy 118 KFSGLNFPNGPVQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCBFKSTYK 177
```

[illegible]

DR	Pfam; PF01353; GFP; 1.
DR	ProDom; PD013756; Green_fl_protein; 1.
SQ	SEQUENCE 227 AA; 26017 MW; 5E312C54EA47F589 CRC64;
Query Match	51.9%; Score 667.5; DB 5; Length 227;
Best Local Similarity	54.9%; Pred. No. 1.5e-54;
Matches 117; Conservative	40; Mismatches 53; Indels 3; Gaps 17;
Qy	1 SVIAKQMTYKYVMGTVNGHYFEVGEVGDKGPKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
Dd	: : : : : : : : : :
Db	2 SVIKPDWKIKLRMEGANVGHKFIEVDGKGKPFEGTQSMDLTVKEGAPLPFPAYDILITVF 61
Qy	61 QYSIGPTKYPEDIPDYVKOSFFGRYTWERIMNMFEDGAVCTVSNDSSIQ---GNCFIYHV 117
Dd	: : : : : : : : : :
Db	62 DYGNRVFAKPQDIPDFVKQTFPEGSWERSMTEYDQGICVATNDITLMKGVDDCFVYKI 121
Qy	118 KFSGLNPPPGPVNMOKKTQGWEPNTERLFARDGMLIGNFMALKLEGGGHVLCBFKSTYK 177
Dd	: : : : : : : : : : :
Db	122 RFDGVNFPANGPVNMOKKTLKWEESTEKMYVRDGVLGKDVNNALLLEGGGHVRCDFKTTYK 181
Qy	178 AKKPVKMPGYHYVDRLDVTNHNKDYTSVEQCE 210
Dd	: : : : : : : : : :
Db	182 AKKFVQLPDHFVDHRIEILLSHDKDNKVKNLYE 214
RESULT 13	
Q7ZOW8	PRELIMINARY; PRF; 227 AA.
ID AC Q7ZOW8	
DT 01-OCT-2003	(TREMBLrel. 25, Created)
DT 01-OCT-2003	(TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003	(TREMBLrel. 25, Last annotation update)
DE	Green fluorescent protein.
OS	Montastraea cavernosa (great star coral).
OC	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
CC	Favilina; Faviidae; Montastraea.
OX	NCBI_TaxID=63558;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=mc2;
EX	MEDLINE=2269801; PubMed=12777529;
RA	Kelmannson I.V., Matz M.V.;
RT	"Molecular Basis and Evolutionary Origins of Color Diversity in Great
RL	Star Coral Montastraea cavernosa (Scleractinia: Faviida).";
RD	Mol. Biol. Evol. 20:1125-1133(2003).
DR	EMBL; AY181553; AAC61599.1; -.
SQ	SEQUENCE 227 AA; 26017 MW; 5E312C54EA47F589 CRC64;
Query Match	51.9%; Score 667.5; DB 5; Length 227;
Best Local Similarity	54.9%; Pred. No. 1.5e-54;
Matches 117; Conservative	40; Mismatches 53; Indels 3; Gaps 17;
Qy	1 SVIAKQMTYKYVMGTVNGHYFEVGEVGDKGPKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
Dd	: : : : : : : : : :
Db	2 SVIKPDWKIKLRMEGANVGHKFIEVDGKGKPFEGTQSMDLTVKEGAPLPFPAYDILITVF 61
Qy	61 QYSIGPTKYPEDIPDYVKOSFFGRYTWERIMNMFEDGAVCTVSNDSSIQ---GNCFIYHV 117
Dd	: : : : : : : : : :
Db	62 DYGNRVFAKPQDIPDFVKQTFPEGSWERSMTEYDQGICVATNDITLMKGVDDCFVYKI 121
Qy	118 KFSGLNPPPGPVNMOKKTQGWEPNTERLFARDGMLIGNFMALKLEGGGHVLCBFKSTYK 177
Dd	: : : : : : : : : : :
Db	122 RFDGVNFPANGPVNMOKKTLKWEESTEKMYVRDGVLGKDVNNALLLEGGGHVRCDFKTTYK 181
Qy	178 AKKPVKMPGYHYVDRLDVTNHNKDYTSVEQCE 210
Dd	: : : : : : : : : :
Db	182 AKKFVQLPDHFVDHRIEILLSHDKDNKVKNLYE 214
RESULT 14	
Q95UA7	PRELIMINARY; PRF; 225 AA.
ID AC Q95UA7	
AC	Q95UA7;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 06:12:47 ; Search time 89.0881 Seconds
(without alignments)
745.314 Million cell updates/sec

Title: US-09-890-463-4

Perfect score: 1287

Sequence: 1 SVIAKQMTYKVMSTGVNGH.....KPVVACRFRRVKSRRHYAVA 235

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 29Jan04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query No.	Score	Match	Length	DB ID	Description
1	1287	100.0	235	3	AAy97150	Pigment p
2	1279	99.4	235	5	ABP70042	Colour Fa
3	1257	97.7	235	5	ABP70026	Colour Fa
4	1242	96.5	231	3	AAy97149	Pigment p
5	1242	96.5	231	5	ABP70025	Colour Fa
6	1190	92.5	226	5	ABP70036	Colour Fa
7	1188	92.3	220	5	ABP70037	Colour Fa
8	1174	91.2	220	5	ABP69926	Colour Fa
9	1174	91.2	223	5	ABP70030	Colour Fa
10	1174	91.2	223	5	ABP70032	Colour Fa
11	1169	90.8	221	5	ABP69992	Colour Fa
12	1169	90.8	221	5	ABP69991	Colour Fa
13	1166	90.6	220	5	ABP70007	Colour Fa
14	1165	90.5	221	5	ABP69967	Colour Fa
15	1165	90.5	221	5	ABP69966	Colour Fa
16	1165	90.5	221	5	ABP70004	Colour Fa
17	1165	90.5	223	5	ABP70033	Colour Fa
18	1165	90.5	235	5	ABP69963	Colour Fa
19	1165	90.5	235	5	ABP69961	Colour Fa
20	1164	90.4	221	5	ABP69978	Colour Fa
21	1163	90.4	223	5	ABP70029	Colour Fa
22	1162	90.3	220	5	ABP69941	Colour Fa
23	1162	90.3	220	5	ABP69940	Colour Fa
24	1161	90.2	220	5	ABP69952	Colour Fa
25	1161	90.2	220	5	ABP69959	Colour Fa

ALIGNMENTS

RESULT 1

AAy97150
ID AAy97150 standard; protein; 235 AA.

XX AC AAy97150;

XX DT 04-DEC-2000 (first entry)

XX DE Pigment protein from coral tissue POC4.

XX N-terminal; pigment protein from coral tissue; PPCT; fluorescence;

KW tissue marker; fluorescent marker; dyestuff; sunscreen; ultra violet;

KW UV filter; POC3.

XX OS Acropora aspera.

XX FH Key Location/Qualifiers

FT Misc-difference 61. .63

FT FT Misc-difference 158 /label= Chromophore_motif

FT FT Misc-difference 192 /note= "critical residue in the vicinity of the fluorophore"

FT FT Misc-difference 192 /note= "critical residue in the vicinity of the fluorophore"

FT FT Misc-difference 210 /note= "critical residue in the vicinity of the fluorophore"

FT FT Misc-difference 210 /note= "critical residue in the vicinity of the fluorophore"

XX WO200046233-A1.

XX PD 10-AUG-2000.

XX 02-FEB-2000; 2000WO-AU0000056.

XX 02-FEB-1999; 99AU-00008463.

XX (UNSY) UNIV SYDNEY.

XX Hoegh-Guldberg O, Dove S;

XX WPI: 2000-532892/48.

XX N-PSDB; AAA52083.

XX Novel pigment protein derived from corals capable of emitting fluorescence upon irradiation by incident light useful as tissue marker, fluorescent marker or general dyestuff.

PS Claim 13; Page 43-44; 49pp; English.

XX cDNA libraries were constructed from a blue pigmented coral, *Acropora*

CC aspera to isolate sequences encoding polypeptides with N-terminal

CC sequences as shown in RAY97147-48. Pigment protein from coral tissue

CC (PPCT) is capable of emitting fluorescence upon irradiation by incident

CC light whose maximal absorbance is in the range of 320-600 nm and a

CC maximal fluorescence emission is in the range of 300-700 nm. PPCT may be

CC used as a tissue marker, fluorescent marker (e.g. to follow gene

CC expression in transformed tissues) or general dyestuff (all claimed).

CC PPCT may also be used in sunscreen formulations or UV filters (both

CC claimed)

XX SQ Sequence 235 AA;

Query Match 100.0%; Score 1287; DB 3; Length 235;

Best Local Similarity 100.0%; Pred. No. 1.2e-127; Length 235;

Matches 235; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVIAKQMTYKVMGTVNGHYFEVEGDKGKPYEGEQTAVLAVTKGGPLPFAWDILSPQC 60

Db 1 SVIAKQMTYKVMGTVNGHYFEVEGDKGKPYEGEQTAVLAVTKGGPLPFAWDILSPQC 60

Qy 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFEDGAVCTVSDSSIQGNCFTYHVKFS 120

Db 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFEDGAVCTVSDSSIQGNCFTYHVKFS 120

Qy 121 GLNFPNGPVNQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCEFKSTYKAKK 180

Db 121 GLNFPNGPVNQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCEFKSTYKAKK 180

Qy 181 PVKMPGYHYVDRKLDVTNNKDYTSVEQCEISIAKRPVACRFRVKSRRHKYAVA 235

Db 181 PVKMPGYHYVDRKLDVTNNKDYTSVEQCEISIAKRPVACRFRVKSRRHKYAVA 235

RESULT 2

ABP70042

ID ABP70042 standard; protein; 235 AA.

XX AC ABP70042;

DT 22-JAN-2003 (first entry)

XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 245.

DE Colour facilitating molecule; CFM; green fluorescent protein; GFP;

XX chromophore; biomatrix; transgenic animal; colouring agent;

XX flower industry; expression marker; reporter molecule; photon trap;

XX UV sink; sunscreen.

OS *Acropora aspera*.

XX WO200270703-A2.

XX 12-SEP-2002.

XX 01-MAR-2002; 2002WO-GB000928.

XX 02-MAR-2001; 2001US-0273227P.

PR 21-MAR-2001; 2001AU-00003874.

PR 15-OCT-2001; 2001US-0329816P.

XX (NUFARM) NUFARM LTD.

PA (UYQU) UNIV QUEENSLAND.

FA (JONE/) JONES E L.

XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;

PI Hoegh-Guldberg IO, Prescott M;

XX WPI; 2002-740765/80.

XX Novel color-facilitating molecule for producing a biomatrix, has a

PT polypeptide which alone/along with molecules imparts altered visual

PT characteristics to cells in the absence of excitation by extraneous non-

XX white light.

XX Example 20; Page 502-503; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)

CC comprising a polypeptide which, in a cell, alone or together with one or

CC more other molecules imparts an altered visual characteristic to the cell

CC when visualised by a human eye in the absence of excitation by extraneous

CC non-white light or particle emission. CFMs are useful for producing a

CC transgenic animal which exhibits a novel colour e.g. sheep with blue or

CC red coloured fleece. They are useful for producing coloured plant

CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other

CC uses include transducing or intensifying an image, providing additional

CC light for growing phototropic organisms e.g. algae and/or corals, for

CC coating materials that experience UV damage e.g. plastics and car

CC upholstery. CFMs are useful in the flower industry, in the development of

CC new varieties of flowering plants. Other contemplated uses include,

CC expression markers, general reporter molecules, photon traps, UV sinks or

CC in sunscreens. CFMs modify visible colour in edible and/or ornamental

CC fungal species, and in fruits and vegetables to enhance their

CC marketability. CFMs embedded in a gel matrix improve image quality in

CC situations of distorted light spectra (biomatrix). The first all-protein

CC chromophore to be isolated was Green Fluorescent protein (GFP). The

CC sequences given in records ABP69924-ABP70048 represent CFM related amino

CC acid sequences

XX SQ Sequence 235 AA;

Query Match 99.4%; Score 1279; DB 5; Length 235;

Best Local Similarity 99.6%; Pred. No. 8.7e-127;

Matches 234; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SVIAKQMTYKVMGTVNGHYFEVEGDKGKPYEGEQTAVLAVTKGGPLPFAWDILSPQC 60

Db 1 SVIAKQMTYKVMGTVNGHYFEVEGDKGKPYEGEQTAVLAVTKGGPLPFAWDILSPQC 60

Qy 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFEDGAVCTVSDSSIQGNCFTYHVKFS 120

Db 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFEDGAVCTVSDSSIQGNCFTYHVKFS 120

Qy 121 GLNFPNGPVNQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCEFKSTYKAKK 180

Db 121 GLNFPNGPVNQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCEFKSTYKAKK 180

Qy 181 PVKMPGYHYVDRKLDVTNNKDYTSVEQCEISIAKRPVACRFRVKSRRHKYAVA 235

Db 181 PVKMPGYHYVDRKLDVTNNKDYTSVEQCEISIAKRPVACRFRVKSRRHKYAVA 235

RESULT 3

ABP70026

ID ABP70026 standard; protein; 235 AA.

XX AC ABP70026;

DT 06-AUG-2003 (revised)

DT 22-JAN-2003 (first entry)

XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 202.

XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;

XX chromophore; biomatrix; transgenic animal; colouring agent;

XX flower industry; expression marker; reporter molecule; photon trap;

XX UV sink; sunscreen.

XX Unidentified.

OS WO200270703-A2.

XX 12-SEP-2002.

```

PF 01-MAR-2002; 2002WO-GB000928.
XX
PR 02-MAR-2001; 2001US-0273227P.
PR 21-MAR-2001; 2001AU-00003874.
PR 15-OCT-2001; 2001US-0329816P.
XX
PA (NUFA-) NUFARM LTD.
PA (UQU) UNIV QUEENSLAND.
PA (JONE/) JONES E L.
XX
PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
PI Hoegh-Guldberg IO, Prescott M;
XX
XX WPI; 2002-740765/80.
XX
XX Novel color-facilitating molecule for producing a biomatrix, has a
XX polypeptide which alone/along with molecules imparts altered visual
XX characteristics to cells in the absence of excitation by extraneous non-
XX white light.
XX
XX Claim 5; Page 479; 510pp; English.
XX
XX The invention relates to an isolated colour-facilitating molecule (CFM)
XX comprising a polypeptide which, in a cell, alone or together with one or
XX more other molecules imparts an altered visual characteristic to the cell
XX when visualised by a human eye in the absence of excitation by extraneous
XX non-white light or particle emission. CFMs are useful for producing a
XX transgenic animal which exhibits a novel colour e.g. sheep with blue or
XX red coloured fleece. They are useful for producing coloured plant
XX extracts, e.g. flavouring, beverage or juice or colouring agent. Other
XX uses include transducing or intensifying an image, providing additional
XX light for growing phototropic organisms e.g. algae and/or corals, for
XX coating materials that experience UV damage e.g. plastics and car
XX upholstery. CFMs are useful in the flower industry, in the development of
XX new varieties of flowering plants. Other contemplated uses include,
XX expression markers, general reporter molecules, photon traps, UV sinks or
XX in sunscreens. CFMs modify visible colour in edible and/or ornamental
XX fungal species, and in fruits and vegetables to enhance their
XX marketability. CFMs embedded in a gel matrix improve image quality in
XX situations of distorted light spectra (biomatrix). The first all-protein
XX chromophore to be isolated was Green Fluorescent protein (GFP). The
XX chromosomes given in records ABP6924-ABP70048 represent CFM related amino
XX acid sequences. (Updated on 06-AUG-2003 to correct OS field.)
XX
XX Sequence 235 AA;
XX
XX Query Match 97.7%; Score 1257; DB 5; Length 235;
XX Best Local Similarity 98.3%; Pred. No. 1.9e-124;
XX Matches 231; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
XX
QY 1 SVIAQMTYKVYMSGTVNGHYFEVGDGKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
DB 1 SVIAQMTYKVYMSGTVNGHYFEVGDGKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
XX
QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFDGAVCTVSDSSIQGNCFIYHVKFS 120
DB 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFDGAVCTVSDSSIQGNCFIYHVKFS 120
XX
QY 121 GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGGHYLCEFKSTYKAKK 180
DB 121 GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGGHYLCEFKSTYKAKK 180
XX
RESULT 4
AAAY97149
XX ID AAY97149 standard; protein; 231 AA.
XX AC
XX AAY97149;
XX

```

```

DT 04-DEC-2000 (first entry)
XX
DE Pigment protein from coral tissue POC3.
XX
KW N-terminal; pigment protein from coral tissue; PPCT; fluorescence;
KW tissue marker; fluorescent marker; dyestuff; sunscreen; ultra violet;
KW UV filter; POC3.
XX
OS Acropora aspera.
XX
XX Key Location/Qualifiers
XX Misc-difference 61..63
XX /label= Chromophore_motif
XX
XX Misc-difference 158
XX /note= "critical residue in the vicinity of the
XX fluorophore"
XX
XX Misc-difference 192
XX /note= "critical residue in the vicinity of the
XX fluorophore"
XX
XX Misc-difference 210
XX /note= "critical residue in the vicinity of the
XX fluorophore"
XX
XX WO200046233-A1.
XX
XX 10-AUG-2000.
XX
XX 02-FEB-2000; 2000WO-AU0000056.
XX
XX 02-FEB-1999; 99AU-00008463.
XX (UNSY) UNIV SYDNEY.
XX
XX Hoegh-Guldberg O, Dove S;
XX WPI; 2000-532892/48.
XX N-PSDB; AAA52082.
XX
XX Novel pigment protein derived from corals capable of emitting
XX fluorescence upon irradiation by incident light useful as tissue marker,
XX fluorescent marker or general dyestuff.
XX
XX Claim 13; Page 42-43; 49pp; English.
XX
XX cDNA libraries were constructed from a blue pigmented coral, Acropora
XX aspera to isolate sequences encoding polypeptides with N-terminal
XX sequences as shown in AAY97147-48. Pigment protein from coral tissue
XX (PPCT) is capable of emitting fluorescence upon irradiation by incident
XX light whose maximal absorbance is in the range of 320-600 nm and a
XX maximal fluorescence emission is in the range of 300-700 nm. PPCT may be
XX used as a tissue marker, fluorescent marker (e.g. to follow gene
XX expression in transformed tissues) or general dyestuff (all claimed).
XX PPCT may also be used in sunscreen formulations or UV filters (both
XX claimed)
XX
XX Sequence 231 AA;
XX
XX Query Match 96.5%; Score 1242; DB 3; Length 231;
XX Best Local Similarity 98.3%; Pred. No. 7e-123;
XX Matches 227; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
XX
QY 1 SVIAQMTYKVYMSGTVNGHYFEVGDGKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
DB 1 SVIAQMTYKVYMSGTVNGHYFEVGDGKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
XX
QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFDGAVCTVSDSSIQGNCFIYHVKFS 120
DB 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFDGAVCTVSDSSIQGNCFIYHVKFS 120
XX
QY 121 GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGGHYLCEFKSTYKAKK 180
DB 121 GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGGHYLCEFKSTYKAKK 180
XX

```

QY 181 PVKMPGYHYVDRKLDVTNNHNDYTSVEQCEISIARKPVVACRFFRVKSRHK 231
|||
Db 181 PVKMPGYHYVDRKLDVTNNHNDYTSVEQCEISIARKPLVACCFRVRKSRHK 231
|||

RESULT 5
ABP70025
ID ABP70025 standard; protein; 231 AA.
XX
AC ABP70025;
XX
DT 06-AUG-2003 (revised)
DT 22-JAN-2003 (first entry)
XX
DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 201.
XX
KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
KW UV sink; sunscreen.
XX
OS Unidentified.
OS
XX
XX WO200270703-A2.
XX
XX 12-SEP-2002.
XX
XX 01-MAR-2002; 2002WO-GH000928.
XX
XX 02-MAR-2001; 2001US-0273227P.
XX
XX 21-MAR-2001; 2001AU-00003874.
XX
XX 15-OCT-2001; 2001US-0329816P.
XX
XX (NUFA-) NUFARM LTD.
XX
XX (UYQU) UNIV QUEENSLAND.
XX
XX (JONE/) JONES E L.
XX
XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
XX
XX Hoegh-Guldberg IO, Prescott M;
XX
XX WPI; 2002-740765/80.
XX
XX Novel color-facilitating molecule for producing a biomatrix, has a
XX polypeptide which alone/along with molecules imparts altered visual
XX characteristics to cells in the absence of excitation by extraneous non-
XX white light.
XX
XX Claim 6; Page 478; 510pp; English.
XX
XX The invention relates to an isolated colour-facilitating molecule (CFM)
XX comprising a polypeptide which, in a cell, alone or together with one or
XX more other molecules imparts an altered visual characteristic to the cell
XX when visualised by a human eye in the absence of excitation by extraneous
XX non-white light or particle emission. CFMs are useful for producing a
XX transgenic animal which exhibits a novel colour e.g. sheep with blue or
XX red coloured fleece. They are useful for producing coloured plant
XX extracts, e.g. flavouring, beverage or juice or colouring agent. Other
XX uses include transducing or intensifying an image, providing additional
XX light for growing phototropic organisms e.g. algae and/or corals, for
XX coating materials that experience UV damage e.g. plastics and car
XX upholstery. CFMs are useful in the flower industry, in the development of
XX new varieties of flowering plants. Other contemplated uses include,
XX expression markers, general reporter molecules, photon traps, UV sinks or
XX in sunscreens. CFMs modify visible colour in edible and/or ornamental
XX fungal species, and in fruits and vegetables to enhance their
XX marketability. CFMs embedded in a gel matrix improve image quality in
XX situations of distorted light spectra (biomatrix). The first all-protein
XX chromophore to be isolated was Green Fluorescent protein (GFP). The
XX sequences given in records ABP6924-ABP70048 represent CFM related amino
XX acid sequences. (Updated on 06-AUG-2003 to correct OS field.)
XX
XX Sequence 231 AA;

Query Match 96.5%; Score 1242; DB 5; Length 231;
Best Local Similarity 98.3%; Pred. No. 7e-123;
Matches 227; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVIAKQMTYKYVMSGTNGHYFEVEGDKGKPYEGEQTIVRLAVTKGGPLPFAWDILSPQC 60
|||
Db 1 SVIAKQMTYKYVMSGTNGHYFEVEGDKGKPYEGEQTIVRLAVTKGGPLPFAWDILSPQC 60
|||

QY 61 QYGSIPFTKYPEDIPDYVKQSPGRTYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKFS 120
|||
Db 61 QYGSIPFTKYPEDIPDYVKQSPGRTYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKFS 120
|||

QY 121 GLNFPNPGVVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGHYLCBFKSTYKAKK 180
|||
Db 121 GLNFPNPGVVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGHYLCBFKSTYKAKK 180
|||

QY 181 PVKMPGYHYVDRKLDVTNNHNDYTSVEQCEISIARKPVVACRFFRVKSRHK 231
|||
Db 181 PVKMPGYHYVDRKLDVTNNHNDYTSVEQCEISIARKPLVACCFRVRKSRHK 231
|||

RESULT 6
ABP70036
ID ABP70036 standard; protein; 226 AA.
XX
XX AC ABP70036;
XX
XX 06-AUG-2003 (revised)
DT 22-JAN-2003 (first entry)
XX
DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 238.
XX
KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
KW UV sink; sunscreen.
XX
OS Unidentified.
OS
XX
XX WO200270703-A2.
XX
XX 12-SEP-2002.
XX
XX 01-MAR-2002; 2002WO-GB000928.
XX
XX 02-MAR-2001; 2001US-0273227P.
XX
XX 21-MAR-2001; 2001AU-00003874.
XX
XX 15-OCT-2001; 2001US-0329816P.
XX
XX (NUFA-) NUFARM LTD.
XX
XX (UYQU) UNIV QUEENSLAND.
XX
XX (JONE/) JONES E L.
XX
XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
XX
XX Hoegh-Guldberg IO, Prescott M;
XX
XX WPI; 2002-740765/80.
XX
XX Novel color-facilitating molecule for producing a biomatrix, has a
XX polypeptide which alone/along with molecules imparts altered visual
XX characteristics to cells in the absence of excitation by extraneous non-
XX white light.
XX
XX Example 19; Page 496-497; 510pp; English.
XX
XX The invention relates to an isolated colour-facilitating molecule (CFM)
XX comprising a polypeptide which, in a cell, alone or together with one or
XX more other molecules imparts an altered visual characteristic to the cell
XX when visualised by a human eye in the absence of excitation by extraneous
XX non-white light or particle emission. CFMs are useful for producing a
XX transgenic animal which exhibits a novel colour e.g. sheep with blue or
XX red coloured fleece. They are useful for producing coloured plant
XX extracts, e.g. flavouring, beverage or juice or colouring agent. Other

CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development of
 CC new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC sequences given in records ABP69924-ABP70048 represent CFM related amino
 CC acid sequences. (Updated on 06-AUG-2003 to correct OS field.)
 XX
 XX Sequence 226 AA;

Query Match 92.5%; Score 1190; DB 5; Length 226;
 Best Local Similarity 96.9%; Pred. No. 2.2e-117;
 Matches 219; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
 QY 1 SVIAKQMTYKVMSTVNGHYFEVGDGKRPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
 DB 1 SVIAKQMTYKVMSTVNGHYFEVGDGKRPYEGEQTVRLAVTKGGPLPFAWDILSPQS 60
 QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKFS 120
 DB 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKFS 120
 QY 121 GLNFPNGPVMQKTKQGWEPNTERLFARDGMLIGNNFALKLEGGHYLCFKSTYKAKK 180
 DB 121 GLNFPNGPVMQKTKQGWEPNTERLFARDGMLIGNNFALKLEGGHYLCFKSTYKAKK 180
 QY 181 PVKMPGYHYVDKLDVTNHNKDYTSVEQCEISIAKPKVPVACFPFRV 226
 DB 181 PVKMPGYHYVDKLDVTNHNKDYTSVEQCEISIAKPKVPVACFPFRV 226

RESULT 7

ABP70037
 ID ABP70037 standard; protein; 220 AA.
 XX
 AC ABP70037;
 XX
 DT 06-AUG-2003 (revised)
 DT 22-JAN-2003 (first entry)
 DE
 DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 239.
 XX
 XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunscreen.
 XX
 OS Unidentified.
 OS
 XX
 PN WO200270703-A2.
 XX
 XX 12-SEP-2002.

01-MAR-2002; 2002WO-GB000928.
 XX
 PR 02-MAR-2001; 2001US-0273227P.
 PR 21-MAR-2001; 2001AU-00003874.
 PR 15-OCT-2001; 2001US-0329816P.
 XX
 XX (NUFA-) NUFARM LTD.
 PA (UYOU) UNIV QUEENSLAND.
 PA (JONE/) JONES E L.
 XX
 XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
 PI Hoeigh-Guldborg IO, Prescott M;
 XX WPI; 2002-740765/80.

XX
 PT Novel color-facilitating molecule for producing a biomatrix, has a
 PT polypeptide which alone/along with molecules imparts altered visual
 PT characteristics to cells in the absence of excitation by extraneous non-
 PT white light.
 XX
 XX Example 19; Page 497-498; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)
 CC comprising a polypeptide which, in a cell, alone or together with one or
 CC more other molecules imparts an altered visual characteristic to the cell
 CC when visualised by a human eye in the absence of excitation by extraneous
 CC non-white light or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant
 CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development of
 CC new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC sequences given in records ABP69924-ABP70048 represent CFM related amino
 CC acid sequences. (Updated on 06-AUG-2003 to correct OS field.)
 XX
 XX Sequence 220 AA;

Query Match 92.3%; Score 1188; DB 5; Length 220;
 Best Local Similarity 98.6%; Pred. No. 3.4e-117;
 Matches 217; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SVIAKQMTYKVMSTVNGHYFEVGDGKRPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
 DB 1 SVIAKQMTYKVMSTVNGHYFEVGDGKRPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
 QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKFS 120
 DB 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKFS 120
 QY 121 GLNFPNGPVMQKTKQGWEPNTERLFARDGMLIGNNFALKLEGGHYLCFKSTYKAKK 180
 DB 121 GLNFPNGPVMQKTKQGWEPNTERLFARDGMLIGNNFALKLEGGHYLCFKSTYKAKK 180
 QY 181 PVKMPGYHYVDKLDVTNHNKDYTSVEQCEISIAKPKVVA 220
 DB 181 PVKMPGYHYVDKLDVTNHNKDYTSVEQCEISIAKPKLVA 220

RESULT 8

ABP69926
 ID ABP69926 standard; protein; 220 AA.

XX
 AC ABP69926;

XX
 DT 22-JAN-2003 (first entry)

XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 24.
 DE
 DE Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunscreen.

XX Acropora aspera.

XX WO200270703-A2.

XX 12-SEP-2002.

```

XX PF 01-MAR-2002; 2002WO-GB000928.
XX PR 02-MAR-2001; 2001US-0273227P.
XX PR 21-MAR-2001; 2001AU-00003874.
XX PR 15-OCT-2001; 2001US-0329816P.
XX PA (NUFA-) NUFARM LTD.
XX PA (UYQU ) UNIV QUEENSLAND.
XX PA (JONE/) JONES E L.
XX PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
XX PI Hoegh-Guldberg IO, Prescott M;
XX DR WPI; 2002-740765/80.
XX PT Novel color-facilitating molecule for producing a biomatrix, has a
XX PT polypeptide which alone/along with molecules imparts altered visual
XX PT characteristics to cells in the absence of excitation by extraneous non-
XX PT white light.
XX PS Claim 5; Page 289; 510pp; English.
XX CC The invention relates to an isolated colour-facilitating molecule (CFM)
XX CC comprising a polypeptide which, in a cell, alone or together with one or
XX CC more other molecules imparts an altered visual characteristic to the cell
XX CC when visualised by a human eye in the absence of excitation by extraneous
XX CC non-white light or particle emission. CFMs are useful for producing a
XX CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
XX CC red coloured fleece. They are useful for producing coloured plant
XX CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
XX CC uses include transducing or intensifying an image, providing additional
XX CC light for growing phototropic organisms e.g. algae and/or corals, for
XX CC coating materials that experience UV damage e.g. plastics and car
XX CC upholstery. CFMs are useful in the flower industry, in the development of
XX CC new varieties of flowering plants. Other contemplated uses include,
XX CC expression markers, general reporter molecules, photon traps, UV sinks or
XX CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
XX CC fungal species, and in fruits and vegetables to enhance their
XX CC marketability. CFMs embedded in a gel matrix improve image quality in
XX CC situations of distorted light spectra (biomatrix). The first all-protein
XX CC chromophore to be isolated was Green Fluorescent protein (GFP). The
XX CC sequences given in records ABP69924-ABP70048 represent CFM related amino
XX CC acid sequences
XX SQ Sequence 220 AA;
XX Query Match 91.2%; Score 1174; DB 5; Length 220;
XX Best Local Similarity 97.7%; Pred. No. 1e-115;
XX Matches 215; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 1 SVIAKQMTYKYVMGTVNGHYFEVEGDKGKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
Db 1 SVIAKQMTYKYVMGTVNGHYFEVEGDKGKPYEGEQTVRLTVTKGGPLPFAWDILSPQS 60
Qy 61 QYGSIPFTKYPEDIPDYVKQSPGRTYRWERIMNFEDGAVCTVSDSISQGNCFIYHVKFS 120
Db 61 QYGSIPFTKYPEDIPDYVKQSPGRTYRWERIMNFEDGAVCTVSDSISQGNCFIYHVKFS 120
Qy 121 GLNFPNPGVNMQKKTQWGPNTERLFARDGMLIGNFMALKEGGHYLCEFKSTYKAKK 180
Db 121 GLNFPNPGVNMQKKTQWGPNTERLFARDGMLIGNFMALKEGGHYLCEFKSTYKAKK 180
Qy 181 PVKMPGHHYVDRKLDVTNNHKNKYTSVEQCEISIAKPKVVA 220
Db 181 PVKMPGHHYVDRKLDVTNNHKNKYTSVEQCEISIAKPKVVA 220
RESULT 9
ABP70030
ID ABP70030 standard; protein; 223 AA.
XX AC
XX AC ABP70030;

```

```

XX DT 22-JAN-2003 (first entry)
XX XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 216.
XX KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
XX KW chromophore; biomatrix; transgenic animal; colouring agent;
XX KW flower industry; expression marker; reporter molecule; photon trap;
XX KW UV sink; sunscreen.
XX OS Tubastrea sp.
XX XX WO200270703-A2.
XX XX 12-SEP-2002.
XX XX 01-MAR-2002; 2002WO-GB000928.
XX XX 02-MAR-2001; 2001US-0273227P.
XX PR 21-MAR-2001; 2001AU-00003874.
XX PR 15-OCT-2001; 2001US-0329816P.
XX XX (NUFA-) NUFARM LTD.
XX PA (UYQU ) UNIV QUEENSLAND.
XX PA (JONE/) JONES E L.
XX PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
XX PI Hoegh-Guldberg IO, Prescott M;
XX DR WPI; 2002-740765/80.
XX XX Novel color-facilitating molecule for producing a biomatrix, has a
XX XX polypeptide which alone/along with molecules imparts altered visual
XX XX characteristics to cells in the absence of excitation by extraneous non-
XX XX white light.
XX PS Example 18; Page 486; 510pp; English.
XX CC The invention relates to an isolated colour-facilitating molecule (CFM)
XX CC comprising a polypeptide which, in a cell, alone or together with one or
XX CC more other molecules imparts an altered visual characteristic to the cell
XX CC when visualised by a human eye in the absence of excitation by extraneous
XX CC non-white light or particle emission. CFMs are useful for producing a
XX CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
XX CC red coloured fleece. They are useful for producing coloured plant
XX CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
XX CC uses include transducing or intensifying an image, providing additional
XX CC light for growing phototropic organisms e.g. algae and/or corals, for
XX CC coating materials that experience UV damage e.g. plastics and car
XX CC upholstery. CFMs are useful in the flower industry, in the development of
XX CC new varieties of flowering plants. Other contemplated uses include,
XX CC expression markers, general reporter molecules, photon traps, UV sinks or
XX CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
XX CC fungal species, and in fruits and vegetables to enhance their
XX CC marketability. CFMs embedded in a gel matrix improve image quality in
XX CC situations of distorted light spectra (biomatrix). The first all-protein
XX CC chromophore to be isolated was Green Fluorescent protein (GFP). The
XX CC sequences given in records ABP69924-ABP70048 represent CFM related amino
XX CC acid sequences
XX SQ Sequence 223 AA;
XX Query Match 91.2%; Score 1174; DB 5; Length 223;
XX Best Local Similarity 97.7%; Pred. No. 1.1e-115;
XX Matches 215; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
Qy 1 SVIAKQMTYKYVMGTVNGHYFEVEGDKGKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
Db 2 SVIAKQMTYKYVMGTVNGHYFEVEGDKGKPYEGEQTVRLTVTKGGPLPFAWDILSPQS 61
Qy 61 QYGSIPFTKYPEDIPDYVKQSPGRTYRWERIMNFEDGAVCTVSDSISQGNCFIYHVKFS 120
Db 62 QYGSIPFTKYPEDIPDYVKQSPGRTYRWERIMNFEDGAVCTVSDSISQGNCFIYHVKFS 121

```

QY 121 GLNPPNGPVVQKKTQGWEPNTERLFARDGMLIGNNFVALKLEGGHYLCBFKSTYKAKK 180
Db 122 GLNPPNGPVVQKKTQGWEPNTERLFARDGMLIGNNFVALKLEGGHYLCBFKSTYKAKK 181
QY 181 PVKMPGYHYVDRKLDVTNHNKDYTSVEQCEISIARKPVVA 220
Db 182 PVKMPGYHYVDRKLDVTNHNKDYTSVEQCEISIARKPVVA 221

RESULT 10

ABP70032
ID ABP70032 standard; protein; 223 AA.

XX AC ABP70032;

XX 22-JAN-2003 (first entry)

XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 220.

XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
KW UV sink; sunscreen.

XX Simularia sp.

XX WO200270703-A2.

XX 12-SEP-2002.

XX 01-MAR-2002; 2002WO-GB000928.

XX 02-MAR-2001; 2001US-0273227P.

XX 21-MAR-2001; 2001AU-00003874.

XX 15-OCT-2001; 2001US-0329816P.

XX (NUFA-) NUFARM LTD.

XX (UYOU) UNIV QUEENSLAND.

XX (JONE/) JONES E L.

XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
PI Hoegh-Guldberg IO, Prescott M;

XX WPI; 2002-740765/80.

XX Novel color-facilitating molecule for producing a biomatrix, has a
PT polypeptide which alone/along with molecules imparts altered visual
PT characteristics to cells in the absence of excitation by extraneous non-
PT white light.

XX Disclosure; Page 489; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)
CC comprising a polypeptide which, in a cell, alone or together with one or
CC more other molecules imparts an altered visual characteristic to the cell
CC when visualised by a human eye in the absence of excitation by extraneous
CC non-white light or particle emission. CFMs are useful for producing a
CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
CC red coloured fleece. They are useful for producing coloured plant
CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
CC uses include transducing or intensifying an image, providing additional
CC light for growing phototropic organisms e.g. algae and/or corals, for
CC coating materials that experience UV damage e.g. plastics and car
CC upholstery. CFMs are useful in the flower industry, in the development of
CC new varieties of flowering plants. Other contemplated uses include,
CC expression markers, general reporter molecules, photon traps, UV sinks or
CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
CC fungal species, and in fruits and vegetables to enhance their
CC marketability. CFMs embedded in a gel matrix improve image quality in
CC situations of distorted light spectra (biomatrix). The first all-protein
CC chromophore to be isolated was Green Fluorescent protein (GFP). The
CC sequences given in records ABP69924-ABP70048 represent CFM related amino

CC acid sequences

XX SQ Sequence 223 AA;

Query Match 91.2%; Score 1174; DB 5; Length 223;

Best Local Similarity 97.7%; Pred. No. 1.1e-115;

Matches 215; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SVIAKQMTYKYVMSCTVNGHYFEVEGDKGKPYEGEOTVRLAVTKGGPLPFAWDILSPQC 60

Db 2 SVIAKQMTYKYVMSCTVNGHYFEVEGDKGKPYEGEOTVRLAVTKGGPLPFAWDILSPQC 61

QY 61 OYGSIPFTKYPEDIPDYVKQSPGRTWERTMNFEDGAVCTVSNDSIQGNCFIYHVKFS 120

Db 62 OYGSIPFTKYLEIDIPDYVKQSPGRTWERTMNFEDGAVCTVSNDSIQGNCFIYHVKFS 121

QY 121 GLNPPNGPVVQKKTQGWEPNTERLFARDGMLIGNNFVALKLEGGHYLCBFKSTYKAKK 180

Db 122 GLNPPNGPVVQKKTQGWEPNTERLFARDGMLIGNNFVALKLEGGHYLCBFKSTYKAKK 181

QY 181 PVKMPGYHYVDRKLDVTNHNKDYTSVEQCEISIARKPVVA 220

Db 182 PVKMPGYHYVDRKLDVTNHNKDYTSVEQCEISIARKPLVA 221

RESULT 11

ABP69992

ID ABP69992 standard; protein; 221 AA.

XX AC ABP69992;

XX 06-AUG-2003 (revised)

XX 22-JAN-2003 (first entry)

XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 149.

XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
KW UV sink; sunscreen.

XX Pocillopora sp.

XX WO200270703-A2.

XX 12-SEP-2002.

XX 01-MAR-2002; 2002WO-GB000928.

XX 02-MAR-2001; 2001US-0273227P.

XX 21-MAR-2001; 2001AU-00003874.

XX 15-OCT-2001; 2001US-0329816P.

XX (NUFA-) NUFARM LTD.

XX (UYOU) UNIV QUEENSLAND.

XX (JONE/) JONES E L.

XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;

PI Hoegh-Guldberg IO, Prescott M;

XX WPI; 2002-740765/80.

XX Novel color-facilitating molecule for producing a biomatrix, has a
PT polypeptide which alone/along with molecules imparts altered visual
PT characteristics to cells in the absence of excitation by extraneous non-
PT white light.

XX Claim 6; Page 435-436; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)
CC comprising a polypeptide which, in a cell, alone or together with one or
CC more other molecules imparts an altered visual characteristic to the cell
CC when visualised by a human eye in the absence of excitation by extraneous

CC non-white light or particle emission. CFMs are useful for producing a
CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
CC red coloured fleece. They are useful for producing coloured plant
CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
CC uses include transducing or intensifying an image, providing additional
CC light for growing phototropic organisms e.g. algae and/or corals, for
CC coating materials that experience UV damage e.g. plastics and car
CC upholstery. CFMs are useful in the flower industry, in the development of
CC new varieties of flowering plants. Other contemplated uses include,
CC expression markers, general reporter molecules, photon traps, UV sinks or
CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
CC fungal species, and in fruits and vegetables to enhance their
CC marketability. CFMs embedded in a gel matrix improve image quality in
CC situations of distorted light spectra (biomatrix). The first all-protein
CC chromophore to be isolated was Green Fluorescent protein (GFP). The
CC sequences given in records ABP69924-ABP70048 represent CFM related amino
CC acid sequences. (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 221 AA;

Query Match 90.8%; Score 1169; DB 5; Length 221;
Best Local Similarity 96.8%; Pred. No. 3.6e-115;
Matches 213; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 SVIAKQMTYKVMSTVNGHYFEVGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
Db 2 SVIATQMTYKVMSTVNGHYFEVGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 61
QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRTWERTWIMNFDGAVCTVSDSSIQGNCFIYHVKFS 120
Db 62 QYGSIPFTKYPEDIPDYVKQSPFGFTWERTWIMNFDGAVCTVSDSSIQGNCFIYHVKFS 121
QY 121 GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGHYLCFFKSTYKAKK 180
Db 122 GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGHYLCFFKSTYKAKK 181
QY 181 PVKMPGYHYVDKLDVTNNKDYTSVEQCEISIARKPVVA 220
Db 182 PVKMPGYHYVDKLDVTNNKDYTSVEQCEISIARKPVVA 221

RESULT 12
ABP69991
ID ABP69991 standard; protein; 221 AA.
XX
AC ABP69991;
DT
DT 06-AUG-2003 (revised)
DT 22-JAN-2003 (first entry)
XX
DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 147.
XX
KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
KW UV sink; sunscreen.

XX Pocillopora sp.
XX
XX WO200270703-A2.
XX
XX 12-SEP-2002.
XX
XX 01-MAR-2002; 2002WO-GB000928.
XX
XX 02-MAR-2001; 2001US-0273227P.
XX 21-MAR-2001; 2001AU-00003874.
XX 15-OCT-2001; 2001US-0329816P.
XX (NUFA-) NUFARM LTD.
XX (UYQU) UNIV QUEENSLAND.
XX (JONE/) JONES E L.

PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
PI Hoegh-Guldberg IO, Prescott M;
XX
DR WPI; 2002-740765/80.
XX
XX Novel color-facilitating molecule for producing a biomatrix, has a
XX polypeptide which alone/along with molecules imparts altered visual
XX characteristics to cells in the absence of excitation by extraneous non-
XX white light.
XX
PS Claim 6; Page 433-434; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)
XX comprising a polypeptide which, in a cell, alone or together with one or
XX more other molecules imparts an altered visual characteristic to the cell
XX when visualised by a human eye in the absence of excitation by extraneous
XX non-white light or particle emission. CFMs are useful for producing a
XX transgenic animal which exhibits a novel colour e.g. sheep with blue or
XX red coloured fleece. They are useful for producing coloured plant
XX extracts, e.g. flavouring, beverage or juice or colouring agent. Other
XX uses include transducing or intensifying an image, providing additional
XX light for growing phototropic organisms e.g. algae and/or corals, for
XX coating materials that experience UV damage e.g. plastics and car
XX upholstery. CFMs are useful in the flower industry, in the development of
XX new varieties of flowering plants. Other contemplated uses include,
XX expression markers, general reporter molecules, photon traps, UV sinks or
XX in sunscreens. CFMs modify visible colour in edible and/or ornamental
XX fungal species, and in fruits and vegetables to enhance their
XX marketability. CFMs embedded in a gel matrix improve image quality in
XX situations of distorted light spectra (biomatrix). The first all-protein
XX chromophore to be isolated was Green Fluorescent protein (GFP). The
XX sequences given in records ABP69924-ABP70048 represent CFM related amino
XX acid sequences. (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 221 AA;

Query Match 90.8%; Score 1169; DB 5; Length 221;
Best Local Similarity 96.8%; Pred. No. 3.6e-115;
Matches 213; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 SVIAKQMTYKVMSTVNGHYFEVGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
Db 2 SVIATQMTYKVMSTVNGHYFEVGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 61
QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRTWERTWIMNFDGAVCTVSDSSIQGNCFIYHVKFS 120
Db 62 QYGSIPFTKYPEDIPDYVKQSPFGFTWERTWIMNFDGAVCTVSDSSIQGNCFIYHVKFS 121
QY 121 GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGHYLCFFKSTYKAKK 180
Db 122 GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGGHYLCFFKSTYKAKK 181
QY 181 PVKMPGYHYVDKLDVTNNKDYTSVEQCEISIARKPVVA 220
Db 182 PVKMPGYHYVDKLDVTNNKDYTSVEQCEISIARKPVVA 221

RESULT 13
ABP70007
ID ABP70007 standard; protein; 220 AA.
XX
XX AC ABP70007;
XX
XX 22-JAN-2003 (first entry)
XX
XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 177.
XX
XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;
XX chromophore; biomatrix; transgenic animal; colouring agent;
XX flower industry; expression marker; reporter molecule; photon trap;
XX UV sink; sunscreen.
XX
XX Montipora sp.

Db 2 SVIATQMTYKYVMSTVNGHYFEVEGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 61
QY 61 QYGSIPFTKYPEDIPDYVKQSPGGRYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKFS 120
Db 62 QYGSIPFTKYPEDIPDYVKQSPGGRYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKFS 121
QY 121 GLNPPNGPVNMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGCHYLCEFKSTYKAKK 180
Db 122 GLNPPNGPVNMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGCHYLCEFKSTYKAKK 181
QY 181 PVKMPGYHYVDKLDVTNNHNDYTSVEQCEISITRKPVVA 220
Db 182 PVKMPGYHYVDKLDVTNNHNDYTSVEQCEISITRKPVVA 221

RESULT 15
ABP69966
ID ABP69966 standard; protein; 221 AA.
XX
AC ABP69966;
XX
DT 06-AUG-2003 (revised)
DT 22-JAN-2003 (first entry)
XX
DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 100.
XX
KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
KW UV sink; sunsreen.
XX
OS Acropora aspera.
XX
PN WO200270703-A2.
XX
PD 12-SEP-2002.
XX
PF 01-MAR-2002; 2002WO-GB000928.
XX
PR 02-MAR-2001; 2001US-0273227P.
PR 21-MAR-2001; 2001AU-00003874.
PR 15-OCT-2001; 2001US-0329816P.
XX
(NUFA-) NUFARM LTD.
PA (UYOU) UNIV QUEENSLAND.
PA (JONE/) JONES E L.
XX
PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
PI Hoegh-Guldberg IO, Prescott M;
XX
WPI; 2002-740765/80.

Novel color-facilitating molecule for producing a biomatrix, has a polypeptide which alone/along with molecules imparts altered visual characteristics to cells in the absence of excitation by extraneous non-white light.

Claim 5; Page 381-382; 510pp; English.

The invention relates to an isolated colour-facilitating molecule (CFM) comprising a polypeptide which, in a cell, alone or together with one or more other molecules imparts an altered visual characteristic to the cell when visualised by a human eye in the absence of excitation by extraneous non-white light or particle emission. CFMs are useful for producing a transgenic animal which exhibits a novel colour e.g. sheep with blue or red coloured fleece. They are useful for producing coloured plant extracts, e.g. flavouring, beverage or juice or colouring agent. Other uses include transducing or intensifying an image, providing additional light for growing phototropic organisms e.g. algae and/or corals, for coating materials that experience UV damage e.g. plastics and car upholstery. CFMs are useful in the flower industry, in the development of new varieties of flowering plants. Other contemplated uses include, expression markers, general reporter molecules, photon traps, UV sinks or

CC in sunscreens. CFMs modify visible colour in edible and/or ornamental fungal species, and in fruits and vegetables to enhance their marketability. CFMs embedded in a gel matrix improve image quality in situations of distorted light spectra (biomatrix). The first all-protein chromophore to be isolated was Green Fluorescent protein (GFP). The sequences given in records ABP69924-ABP70048 represent CFM related amino acid sequences. (Updated on 06-AUG-2003 to correct OS field.)
XX
SQ Sequence 221 AA;

Query Match 90.5%; Score 1165; DB 5; Length 221;
Best Local Similarity 96.4%; Pred. No. 9.5e-115;
Matches 212; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 SVIATQMTYKYVMSTVNGHYFEVEGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
Db 2 SVIATQMTYKYVMSTVNGHYFEVEGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 61
QY 61 QYGSIPFTKYPEDIPDYVKQSPGGRYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKFS 120
Db 62 QYGSIPFTKYPEDIPDYVKQSPGGRYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKFS 121
QY 121 GLNPPNGPVNMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGCHYLCEFKSTYKAKK 180
Db 122 GLNPPNGPVNMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGCHYLCEFKSTYKAKK 181
QY 181 PVKMPGYHYVDKLDVTNNHNDYTSVEQCEISITRKPVVA 220
Db 182 PVKMPGYHYVDKLDVTNNHNDYTSVEQCEISITRKPVVA 221

Search completed: August 12, 2004, 06:17:06
Job time : 90.0881 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 06:19:43 ; Search time 494.559 Seconds
(without alignments)
149.169 Million cell updates/sec

Title: US-09-890-463-4
Perfect score: 1287
Sequence: 1 SVIAKQMTYKVMSTGVNGH.....KPVVACRRFRVKSRYKAVA 235

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	788	61.2	225	15	US-10-442-148A-7
2	788	61.2	239	15	US-10-442-148A-8
3	785	61.0	225	14	US-10-315-920-6
4	784	60.9	225	14	US-10-121-258-20
5	784	60.9	225	14	US-10-315-920-4
6	784	60.9	236	16	US-10-314-936-2
7	784	60.8	236	16	US-10-314-936-4
8	783	60.8	225	9	US-09-999-745-67
9	783	60.8	225	10	US-09-866-538-12
10	783	60.8	225	10	US-09-794-308-12
11	783	60.8	225	10	US-09-865-231-12
12	783	60.8	225	12	US-10-132-067-4
13	783	60.8	225	13	US-10-006-922-12
14	783	60.8	225	14	US-10-081-864-8
15	783	60.8	225	14	US-10-121-258-1

16	783	60.8	225	14	US-10-315-920-2
17	783	60.8	225	15	US-10-370-570-56
18	783	60.8	225	15	US-10-406-618-32
19	783	60.8	225	16	US-10-433-640-13
20	783	60.8	240	14	US-10-152-296-2
21	783	60.8	240	16	US-10-739-656-2
22	783	60.8	487	15	US-10-343-977-1
23	783	60.8	505	15	US-10-343-977-2
24	783	60.8	545	14	US-10-214-932-52
25	783	60.8	547	15	US-10-343-977-3
26	783	60.8	548	14	US-10-214-932-76
27	778	60.5	225	16	US-10-423-688A-41
28	773	60.1	227	13	US-10-006-922-10
29	769	59.8	225	14	US-10-121-258-24
30	769	59.8	226	14	US-10-121-258-6
31	768	59.7	225	13	US-10-006-922-44
32	768	59.7	225	14	US-10-081-864-12
33	768	59.7	225	14	US-10-121-258-4
34	764.5	59.4	226	16	US-10-724-178-12
35	750	58.3	230	13	US-10-006-922-18
36	750	58.3	230	14	US-10-161-403-40
37	737.5	57.3	232	14	US-10-133-973-5
38	737.5	57.3	232	15	US-10-370-570-64
39	729.5	56.7	232	16	US-10-724-178-11
40	717	55.7	225	14	US-10-121-258-8
41	717	55.7	225	16	US-10-724-178-16
42	706	54.9	225	15	US-10-370-570-61
43	689	53.5	205	13	US-10-006-922-46
44	671	52.1	231	14	US-10-133-973-6
45	671	52.1	266	13	US-10-006-922-4

ALIGNMENTS

RESULT 1
US-10-442-148A-7
; Sequence 7, Application US/10442148A
; Publication No. US20040014242A1
; GENERAL INFORMATION:
; APPLICANT: IWAKURA, KASAHIRO
; APPLICANT: HIROTA, KIYONORI
; TITLE OF INVENTION: PROCESS FOR IMMOBILIZING ORIENTATION-CONTROLLED PROTEIN AND
; FILE REFERENCE: 04583.0103-00000
; CURRENT APPLICATION NUMBER: US/10/442,148A
; CURRENT FILING DATE: 2003-05-21
; PRIOR APPLICATION NUMBER: JP 2002-148950
; PRIOR FILING DATE: 2002-05-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic protein sequence
US-10-442-148A-7

Query Match	61.2%	Score 788;	DB 15;	Length 225;
Best Local Similarity	63.0%;	Pred. No. 4.3e-76;		
Matches 145;	Conservative 27;	Mismatches 44;	Indels 14;	Gaps 1;
QY	1	SVIAKQMTYKVMSTGVNGHYFEVGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC	60	
Db	6	NVITEFMRKPMVEGTGHEFEIEGEGRPYKLVKTKVTKGGPLPFAWDILSPQF	65	
QY	61	QVGSFPFTKPEDIDPVYKQSPFGYTWERTMNRDGAQCTVSDSSIQNCFIYHVKFS	120	
Db	66	QVGSKVVYKHADIDPVYKLSFPFGKWRVWNFEDGGVATVQSSLDGCFYKVFYI	125	
QY	121	GLNFPFGVPMOKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCFKSTYKARK	180	

```
Db 126 GVNFPDGVNMKGKTMGWEASTERLYPRDGLKGETHKALKLDGGHYLVBFKSIYMAKK 185
Qy 181 PVKMPGHHYVDRKLDVTNNHNDYTSVQCEISIAKPPVACRFFRVKSRH 230
Db 186 PVQLPGYIYVDAKLDITSHNEDYIVEQYE-----RTEGRH 221

RESULT 2
US-10-442-148A-8
; Sequence 8, Application US/10442148A
; Publication No. US20040014242A1
; GENERAL INFORMATION:
; APPLICANT: IWAKURA, MASAHIRO
; APPLICANT: HIROTA, KIYONORI
; TITLE OF INVENTION: PROCESS FOR IMMOBILIZING ORIENTATION-CONTROLLED PROTEIN AND
; FILE REFERENCE: 04583.0103-00000
; CURRENT APPLICATION NUMBER: US/10/442,148A
; CURRENT FILING DATE: 2003-05-21
; PRIOR APPLICATION NUMBER: JP 2002-148950
; PRIOR FILING DATE: 2002-05-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic protein sequence
US-10-442-148A-8

Query Match 61.2%; Score 788; DB 15; Length 239;
Best Local Similarity 63.0%; Pred. No. 4.7e-76;
Matches 145; Conservative 27; Mismatches 44; Indels 14; Gaps 1;

Qy 1 SVIAKQMTYKYVMSGTVNGHYFVEVGDKGKPYEGEQTIVRLAVTKGGPLPFAWDILSPQC 60
Db 6 NVITEFMRFKVMEGTVNGHFEFEIGEGERPYEGHNTVKLVTKGGPLPFAWDILSPQF 65
Qy 61 QYGSIPFTKYPEDIPDYVKQSPGGRYTWERIMNPFEDGAVCTVSDSSIQGNCFTYHVKFS 120
Db 66 QYGSKVYVKGHPADIPDYKLSFPFGFKWERVMNPFEDGGVAIVTQDSSLQDGCFTYKVKFI 125
Qy 121 GLNFPNGPVNMKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCBFKSTYKAKK 180
Db 126 GVNFPDGVNMKGKTMGWEASTERLYPRDGLKGETHKALKLDGGHYLVBFKSIYMAKK 185
Qy 181 PVKMPGHHYVDRKLDVTNNHNDYTSVQCEISIAKPPVACRFFRVKSRH 230
Db 186 PVQLPGYIYVDAKLDITSHNEDYIVEQYE-----RTEGRH 221

RESULT 3
US-10-315-920-6
; Sequence 6, Application US/10315920
; Publication No. US20030175809A1
; GENERAL INFORMATION:
; APPLICANT: Fradkov, Arcady Fedorovich
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
; FILE REFERENCE: CLON-077CIP
; CURRENT APPLICATION NUMBER: US/10/315,920
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/211,607
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: PCT/US01/19097
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DeRed with I125R
US-10-121-258-20

Query Match 60.9%; Score 784; DB 14; Length 225;
Best Local Similarity 63.0%; Pred. No. 1.2e-75;
Matches 145; Conservative 27; Mismatches 44; Indels 14; Gaps 1;

Qy 1 SVIAKQMTYKYVMSGTVNGHYFVEVGDKGKPYEGEQTIVRLAVTKGGPLPFAWDILSPQC 60
Db 6 NVIKEFMRFKVMEGTVNGHFEFEIGEGERPYEGHNTVKLVTKGGPLPFAWDILSPQF 65
Qy 61 QYGSIPFTKYPEDIPDYVKQSPGGRYTWERIMNPFEDGAVCTVSDSSIQGNCFTYHVKFS 120
Db 66 QYGSKVYVKGHPADIPDYKLSFPFGFKWERVMNPFEDGGVAIVTQDSSLQDGCFTYKVKFR 125
Qy 121 GLNFPNGPVNMKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCBFKSTYKAKK 180
Db 126 GVNFPDGVNMKKTMGWEASTERLYPRDGLKGETHKALKLDGGHYLVBFKSIYMAKK 185
Qy 181 PVKMPGHHYVDRKLDVTNNHNDYTSVQCEISIAKPPVACRFFRVKSRH 230
Db 186 PVQLPGYIYVDAKLDITSHNEDYIVEQYE-----RTEGRH 221
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: variant of sequence from Discosoma sp.
US-10-315-920-6

Query Match 61.0%; Score 785; DB 14; Length 225;
Best Local Similarity 63.0%; Pred. No. 9.1e-76;
Matches 145; Conservative 27; Mismatches 44; Indels 14; Gaps 1;

Qy 1 SVIAKQMTYKYVMSGTVNGHYFVEVGDKGKPYEGEQTIVRLAVTKGGPLPFAWDILSPQC 60
Db 6 NVITEFMRFKVMEGTVNGHFEFEIGEGERPYEGHNTVKLVTKGGPLPFAWDILSPQF 65
Qy 61 QYGSIPFTKYPEDIPDYVKQSPGGRYTWERIMNPFEDGAVCTVSDSSIQGNCFTYHVKFS 120
Db 66 QYGSKVYVKGHPADIPDYKLSFPFGFKWERVMNPFEDGGVAIVTQDSSLQDGCFTYKVKFI 125
Qy 121 GLNFPNGPVNMKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCBFKSTYKAKK 180
Db 126 GVNFPDGVNMKKTMGWEASTERLYPRDGLKGETHKALKLDGGHYLVBFKSIYMAKK 185
Qy 181 PVKMPGHHYVDRKLDVTNNHNDYTSVQCEISIAKPPVACRFFRVKSRH 230
Db 186 PVQLPGYIYVDAKLDITSHNEDYIVEQYE-----RTEGRH 221

RESULT 4
US-10-121-258-20
; Sequence 20, Application US/10121258
; Publication No. US20030059835A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; FILE REFERENCE: UC083.1C2CPI
; CURRENT APPLICATION NUMBER: US/10/121,258
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DeRed with I125R
US-10-121-258-20

Query Match 60.9%; Score 784; DB 14; Length 225;
Best Local Similarity 63.0%; Pred. No. 1.2e-75;
Matches 145; Conservative 27; Mismatches 44; Indels 14; Gaps 1;

Qy 1 SVIAKQMTYKYVMSGTVNGHYFVEVGDKGKPYEGEQTIVRLAVTKGGPLPFAWDILSPQC 60
Db 6 NVIKEFMRFKVMEGTVNGHFEFEIGEGERPYEGHNTVKLVTKGGPLPFAWDILSPQF 65
Qy 61 QYGSIPFTKYPEDIPDYVKQSPGGRYTWERIMNPFEDGAVCTVSDSSIQGNCFTYHVKFS 120
Db 66 QYGSKVYVKGHPADIPDYKLSFPFGFKWERVMNPFEDGGVAIVTQDSSLQDGCFTYKVKFR 125
Qy 121 GLNFPNGPVNMKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCBFKSTYKAKK 180
Db 126 GVNFPDGVNMKKTMGWEASTERLYPRDGLKGETHKALKLDGGHYLVBFKSIYMAKK 185
Qy 181 PVKMPGHHYVDRKLDVTNNHNDYTSVQCEISIAKPPVACRFFRVKSRH 230
Db 186 PVQLPGYIYVDAKLDITSHNEDYIVEQYE-----RTEGRH 221
```

QY	1	SVIAKQMTYKVMGCTVNGHYFEVGEQVRLAVTKGGLPFAWDILSPQC	60
Db	6	NVIKFMRFKVRMEGTVNGHFEIEGEGRPPYGHNTVKLVKTKGGLPFAWDILSPQF	65
QY	61	QYGSIPFTKYPEDIPDYVKQSPGRTYMERIMNFEDGAVCTVSDSSIQGNCFTYHVKFS	120
Db	66	QYGSKVYVHPADIPDYKLSFPEGFKWERVMNFEDGGVTVTQDSLQDGCFTYKVKFI	125
QY	121	GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNFMALKEGGHYLCFEKSTYKAKK	180
Db	126	GVNFPDGPVMQKKTGMWEASTERLYPRDGLVKGELHAKLKDGGHYLVFEKTYMAKK	185
QY	181	PVKMPGYHYVDRKLDVTNNHNDYTSVEQCEISIARKPVVACRFFRVKSRH	230
Db	186	PVQLPGYYVDSKLDITSHNKDYTIIVEQYE-----RTEGRH	221
RESULT 7			
US-10-314-936-4			
; Sequence 4, Application US/10314936			
; Publication No. US20040110225A1			
; GENERAL INFORMATION:			
; APPLICANT: Gibbs, Patrick D.L.			
; APPLICANT: Carter, Robert W.			
; APPLICANT: Schmale, Michael C.			
; TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES			
; FILE REFERENCE: 638.004			
; CURRENT APPLICATION NUMBER: US/10/314,936			
; CURRENT FILING DATE: 2002-12-09			
; NUMBER OF SEQ ID NOS: 15			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 4			
; LENGTH: 236			
; TYPE: PRT			
; ORGANISM: Artificial			
; FEATURE:			
; OTHER INFORMATION: mutant red fluorescent protein			
US-10-314-936-4			
Query Match 60.9%; Score 784; DB 16; Length 236;			
Best Local Similarity 63.0%; Pred. No. 1.2e-75;			
Matches 145; Conservative 27; Mismatches 44; Indels 14; Gaps 1;			
QY	1	SVIAKQMTYKVMGCTVNGHYFEVGEQVRLAVTKGGLPFAWDILSPQC	60
Db	6	NVIKFMRFKVRMEGTVNGHFEIEGEGRPPYGHNTVKLVKTKGGLPFAWDILSPQF	65
QY	61	QYGSIPFTKYPEDIPDYVKQSPGRTYMERIMNFEDGAVCTVSDSSIQGNCFTYHVKFS	120
Db	66	QYGSKVYVHPADIPDYKLSFPEGFKWERVMNFEDGGVTVTQDSLQDGCFTYKVKFI	125
QY	121	GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNFMALKEGGHYLCFEKSTYKAKK	180
Db	126	GVNFPDGPVMQKKTGMWEASTERLYPRDGLVKGELHAKLKDGGHYLVFEKTYMAKK	185
QY	181	PVKMPGYHYVDRKLDVTNNHNDYTSVEQCEISIARKPVVACRFFRVKSRH	230
Db	186	PVQLPGYYVDSKLDITSHNKDYTIIVEQYE-----RTEGRH	221
RESULT 8			
US-09-999-745-67			
; Sequence 67, Application US/09999745			
; Patent No. US20020157120A1			
; GENERAL INFORMATION:			
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA			
; APPLICANT: Tsien, Roger Y.			
; APPLICANT: Baird, Geoffrey			
; TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS			
; FILE REFERENCE: REGEN1470-1			
; CURRENT APPLICATION NUMBER: US/09/999,745			
; CURRENT FILING DATE: 2001-10-23			
; PRIOR APPLICATION NUMBER: 09/316,920			

```

; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 67
; LENGTH: 325
; TYPE: PR1
; ORGANISM: Discosoma sp.
US-09-999-745-67

```

[illegible]

-RESULT 9

```

US-09-866-538-12
; Sequence 12, Application US/09866538
; Publication No. US20030032088A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TS'EN, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
; FILE REFERENCE: REGEN1530-2
; CURRENT APPLICATION NUMBER: US/09/866,538
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-09-866-538-12

```

Query Match	60.8%;	Score 783;	DB 10;	Length 225;
Best Local Similarity	63.0%;	Pred. No. 1.5e-75;		
Matches 145; Conservative 27; Mismatches 44; Indels 14; Gaps 1;				
QY	1	SVIAKQMTVKVMSGTWNGHYFVEVGDGKPYEGEQTVRLAVTKGGPLPFPAMDILSPQC	60	
DB	6	NVIKEPWRFKVRMEGTWNGHEFLEBEGEGRPTEGHTVVKLVTKGGPLPFPAMDILSPQF	65	
QY	61	QYGSIBFTKYPEDIPYVQKSPFGRYTWERIMNFDGAVCTVSNDDSSIQNCFTFYHYKFS	120	
DB	66	QYGSKYVYVHPADIPDYKLSFEPGFKWERNVNFDDGGVYVTVDSSLDQGCFTYKVKFI	125	
QY	121	GLNFPPNGPVWQKKTQGWEPNTERLFDARDQMLIGNNFMALKLEGGGHYLCBPKSTYKAKK	180	
DB	126	GVNFPDGPVWQKKTGWGEASTERLPDRDGLVKEITHKALKLDGGHYLVEFKSIYMAKK	185	
QY	181	PVKMPGHYVDRKLDVTNNHKNQVTSVEQCEISIAKPKPVVACRFFRVKSRH	230	
DB	186	PVLGPYVYVDSKLDITSHNEDITYVEQYE-----RTGRH	221	

RESULT 10

US-09-794-308-12

```

; Sequence 12, Application US/09794308
; Publication No. US20030170911A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TS'EN, Roger
; APPLICANT: ZACHARIAS, David
; APPLICANT: BAIRD, Geoffrey
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
; FILE REFERENCE: REGEN1530
; CURRENT APPLICATION NUMBER: US/09/794,308
; CURRENT FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
; US-09-794-308-12

```

Query Match	60.8%;	Score	783;	DB	10;	Length	225;		
Best Local Similarity	63.0%;	Pred. No.	1.5e-75;						
Matches	145;	Conservative	27;	Mismatches	44;	Indels	14;	Gaps	17;
QY	1	SVIAKOWTVKVMASGTVNGHYFVEVGDKPKVEGEOTVRLAVTKGGLPFPAMDILSPQC	60						
DB	6	NVLKEFWRFKVRMEGTVNGHEFIEGEGEGRPEGHTVKLVTKGGLPFPAMDILSPQF	65						
QY	61	QYGSIPFTKYPEDIPDYVKOSFPGRYTWRIMMFEDGAVCTVSNDDSIQGNCFITYHWKFS	120						
DB	66	QYGSKVYVHKPADIPTYKLSFPEGFKWERVMNFEDGGVTVTODSSLQDGCFTYKVKFI	125						
QY	121	GLNFPPNGPMQKKTQGWEPNTERLFARDQMLIGNNFMALKLEGGHVLCFEKSTYKAKK	180						
DB	126	GVNFPSDGPMQKKTGMCWEASTERLPYPRDGVLGKEIHKALKLKDGGHYLVFEKSIYKAKK	185						
QY	181	PVKMPGVHYVDRLKLDVTNNHKDYTSVEQCEISIAARKPVPVACRFPFVKSRRH	230						
DB	186	PVLGPQYVYVDSKLDITSHNEDYTVVEQYE-----RTGRH	221						

RESULT 11

```

US-09-865-291-12
; Sequence 12, Application US/09865291
; Publication No. US20030186229A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF CALIFORNIA
; APPLICANT: TS'EN, Roger
; APPLICANT: T'ING, Alice
; APPLICANT: ZHANG, Jin
; TITLE OF INVENTION: EMISSION RATIO-METRIC INDICATORS OF PHOSPHORYLATION
; FILE REFERENCE: REGEN1550
; CURRENT APPLICATION NUMBER: US/09/865,291
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-09-865-291-12

```

Query Match 60.8%; Score 783; DB 10; Length 225;
Best Local Similarity 63.0%; Pred. No. 1.5e-75;
Matches 145; Conservative 27; Mismatches 44; Indels 14; Gaps 1;

```
QY      1 SVIAKMTKKVMSGTNGHYFVEVGDKGPKVEGEQTVELATVKGGPLPFANDILSPQC 60  
       :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|  
Db      6 NVIKFPMKVRMEGTNGHGFIEEGEGEPYEGHNTVKLVKTGGPLPFANDILSPQF 65  
  
QY     61 QYGSIFPTKYPIIDPYIVKQSPPGRYTWERIMNFDGAVCTVGNDSIQNCFTIYHKFS 120  
       ||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db     66 QYGSKVYVHKPADIPIYKKLSPFGFKWRVMNFEDGVVTVTQDSLSDGCFTYKVKFI 125
```


D _b	66	QYGSVVYKHPADIPDYKKLSPBGFKWERNVNFEDGGVVTVTQDSSLDGQCFIYVKFI	125
Q _Y	121	GLNPPNGPVNKKQTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCEPKSTYKAKK	180
D _b	126	GVNFPDGPVVKKTMGWEASTERLYPRDGVLKGELHKALKDKGGHYLVEPKSYMAKK	185
Q _Y	181	PVKMPGYHYVDRKLDVTVNNKDYTSVEQCEISIAKKPVVACFRFVKSRH	230
D _b	186	PVQLPGYYVDPKSLDITSHNEDYTVVEQVE-----RTEGSH	221

```

RESULT 15
US-10-121-258-1
; Sequence 1, Application US/10121258
; Publication No. US20030059835A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; FILE OF INVENTION: PROTEIN VARIANTS AND METHODS FOR MAKING SAME
; TITLE REFERENCE: UC083.1CP2CP1
; CURRENT APPLICATION NUMBER: US/10/121,258
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(225)
; OTHER INFORMATION: wild-type DsRed
US-10-121-258-1

```

	Query Match	60.8%;	Score	783;	DB	14;	Length	225;
	Best Local Similarity	63.0%;	Pred.	No. 1.5e-75;				
	Matches	145;	Conservative	27;	Mismatches	44;	Indels	14;
	Gaps							
Qy	1 SVIAKQMITYKYVMGSGTVNGHYFVEVGDKGKPYEGETVRLAVTKGGPLPFAWDILSPQC	60						
	: : : : : : : : : : : : : : : : : : :							
Db	6 NVIKEFWFKVRMGSTVNGHFEFEIEGBEGRPYEGHTVKLVKTGGGLPFAWDILSPQF	65						
	: : : : : : : : : : : : : : : : : : :							
Qy	61 QYGSIPFTKPEDIPDVVKOSFPGRYTWERIMNFDGACVTVSNDSSIQGNCFITYHKFS	120						
	: : : : : : : : : : : : : : : : : : :							
Db	66 QYGSKVYVKHPADIPDYKKLSFPFGFKWERYMNFEDGGVVTVTQDSSLQDCFTYKVKFI	125						
	: : : : : : : : : : : : : : : : : : :							
Qy	121 GLNPFPNGPVMKKTQGWEPNTELFARDGMLIGNFMALKEGGHYLCSFKETYKAKK	180						
	: : : : : : : : : : : : : : : : : : :							
Db	126 GVNFPSDPGVMKKTGMGEASTERLYPRDGVLKCEIHKALKKGUGHYLVEFKSIYWAKK	185						
	: : : : : : : : : : : : : : : : : : :							
Qy	181 PVKMFGHYVDRLKDVTNNKHNDYTSVBGCETSIARKPVVACRPFVRSRH	230						
	: : : : : : : : : : : : : : : : : : :							
Db	186 PQVLPGYYVDSKLDTITSHNEDYTIIVBQE-----RTGSRH	221						

Search completed: August 12, 2004, 06:51:22
Job time : 495.559 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 07:02:25 ; Search time 40 Seconds
(without alignments)
21.941 Million cell updates/sec

Title: US-09-890-463-2

Perfect score: 83

Sequence: 1 SVIAKQMTYKYVMSGTV 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 157007

Minimum DB seq length: 0

Maximum DB seq length: 17

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	29	34.9	14	1	US-08-443-104-2
2	29	34.9	14	1	US-08-238-130-3
3	29	34.9	14	1	US-08-442-859-2
4	29	34.9	14	2	US-08-398-489-2
5	29	34.9	14	2	US-08-894-772-3
6	29	34.9	14	2	US-09-207-844-3
7	29	34.9	14	5	PCT-US95-05534-2
8	28	33.7	12	4	US-09-528-200-40
9	28	33.7	13	2	US-08-413-708B-6
10	28	33.7	13	4	US-09-528-200-33
11	28	33.7	13	4	US-09-528-200-39
12	28	33.7	14	4	US-09-528-200-26
13	28	33.7	14	4	US-09-528-200-32
14	28	33.7	15	1	US-08-486-721A-16
15	28	33.7	15	4	US-09-528-200-19
16	28	33.7	15	4	US-09-528-200-25
17	28	33.7	16	4	US-09-528-200-12
18	28	33.7	16	4	US-09-528-200-18
19	28	33.7	17	1	US-07-924-054-7
20	28	33.7	17	4	US-09-528-200-11
21	27	32.5	10	4	US-09-528-200-42
22	27	32.5	11	4	US-09-528-200-35
23	27	32.5	11	4	US-09-528-200-41
24	27	32.5	11	4	US-09-528-200-47
25	27	32.5	11	4	US-09-528-200-190
26	27	32.5	11	4	US-09-528-200-192
27	27	32.5	12	3	US-08-512-560-7

28	27	32.5	12	4	US-09-528-200-28	Sequence 28, Appl
29	27	32.5	12	4	US-09-528-200-34	Sequence 34, Appl
30	27	32.5	12	4	US-09-528-200-46	Sequence 46, Appl
31	27	32.5	12	4	US-09-528-200-191	Sequence 191, Appl
32	27	32.5	13	3	US-08-952-568-21	Sequence 21, Appl
33	27	32.5	13	4	US-09-246-963A-24	Sequence 24, Appl
34	27	32.5	13	4	US-09-528-200-21	Sequence 21, Appl
35	27	32.5	13	4	US-09-528-200-27	Sequence 27, Appl
36	27	32.5	14	3	US-08-930-845-6	Sequence 6, Appl
37	27	32.5	14	3	US-08-952-568-23	Sequence 23, Appl
38	27	32.5	14	3	US-09-192-048-25	Sequence 25, Appl
39	27	32.5	14	4	US-09-528-200-14	Sequence 14, Appl
40	27	32.5	14	4	US-09-528-200-20	Sequence 20, Appl
41	27	32.5	14	6	5194585-13	Patent No. 5194585
42	27	32.5	15	4	US-09-528-200-13	Sequence 13, Appl
43	27	32.5	17	3	US-08-896-933-12	Sequence 12, Appl
44	27	32.5	17	3	US-09-288-391-24	Sequence 24, Appl
45	27	32.5	17	4	US-09-314-235-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-443-104-2
; Sequence 2, Application US/08443104
; Patent No. 5691162
; GENERAL INFORMATION:
; APPLICANT: Shuster, Jeffrey R.
; APPLICANT: Madden, Mark
; APPLICANT: Moyer, Donna L.
; APPLICANT: Fuglsang, Claus
; APPLICANT: Branner, Sven
; TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING INCREASED
; TITLE OF INVENTION: ACTIVITY
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSER: No. 56911620 No. 5691162disk of No. 5691162th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/443,104
; FILING DATE: 17-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/398,489
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Agis Dr. Cheryl H.
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 4180.010-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-443-104-2

Query Match 34.9%; Score 29; DB 1; Length 14;
Best Local Similarity 75.0%; Pred. No. 42;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 TYKVYMSG 15
|||||
Db 2 TYKVYPWG 9

RESULT 2

US-08-238-130-3
; Sequence 3, Application US/08238130
; Patent No. 5702934
; GENERAL INFORMATION:
; APPLICANT: Hastrup, Sven
; APPLICANT: Branner, Sven
; APPLICANT: Jorgensen, Birthe R.
; APPLICANT: Christensen, Tove
; APPLICANT: Jorgensen, Birgitte B.
; APPLICANT: Shuster, Jeffrey R.
; APPLICANT: Madden, Mark
; APPLICANT: Moyer, Donna L.
; APPLICANT: Fuglsang, Claus
; TITLE OF INVENTION: PROCESSES FOR PRODUCING AN ENZYME
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5702934o No. 5702934disk of No. 5702934th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/238,130
; FILING DATE: 04-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DK 522/93
; FILING DATE: 05-MAY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Agnis Dr., Cheryl H.
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 3965.200-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-238-130-3

Query Match 34.9%; Score 29; DB 1; Length 14;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 TYKVYMSG 15
|||||
Db 2 TYKVYPWG 9

RESULT 3

US-08-442-859-2
; Sequence 2, Application US/08442859
; Patent No. 5807729
; GENERAL INFORMATION:
; APPLICANT: Shuster, Jeffrey R.

; APPLICANT: Madden, Mark
; APPLICANT: Moyer, Donna L.
; APPLICANT: Fuglsang, Claus
; APPLICANT: Branner, Sven
; TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING INCREASED
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5807729o No. 5807729disk of No. 5807729th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/442,859
; FILING DATE: 17-MAY-1995
; CLASSIFICATION: 252
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/398,489
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Agnis Dr., Cheryl H.
; REGISTRATION NUMBER: 34,086
; REFERENCE/DOCKET NUMBER: 4180.010-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-442-859-2

Query Match 34.9%; Score 29; DB 1; Length 14;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 TYKVYMSG 15
|||||
Db 2 TYKVYPWG 9

RESULT 4

US-08-398-489-2
; Sequence 2, Application US/08398489
; Patent No. 5843753
; GENERAL INFORMATION:
; APPLICANT: Shuster, Jeffrey R.
; APPLICANT: Madden, Mark
; APPLICANT: Moyer, Donna L.
; APPLICANT: Fuglsang, Claus
; APPLICANT: Branner, Sven
; TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING INCREASED
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5843753o No. 5843753disk of No. 5843753th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Tape

Query Match 34.9%; Score 29; DB 1; Length 14;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 TYKVYMSG 15
|||||
Db 2 TYKVYPWG 9

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/398,489
FILING DATE: 03-MAR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Agilis Dr., Cheryl H.
REGISTRATION NUMBER: 34,086
REFERENCE/DOCKET NUMBER: 4180.010-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-398-489-2

Query Match 34.9%; Score 29; DB 2; Length 14;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 TYKVYMSG 15
| | | | |
Db 2 TYKVYPWG 9

RESULT 5
US-08-894-772-3
Sequence 3, Application US/08894772
Patent No. 5861280
GENERAL INFORMATION:
APPLICANT: Lehmbeck, Jan
TITLE OF INVENTION: Host Cell Expressing Reduced Levels
TITLE OF INVENTION: Of A Metalloprotease And Methods Using The Host Cell In Protei
TITLE OF INVENTION: Production
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5861280o No. 5861280disk of No. 5861280th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/894,772
FILING DATE: 27-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeta A.
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 4300.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-867-0298
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-894-772-3

Query Match 34.9%; Score 29; DB 2; Length 14;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 TYKVYMSG 15
| | | | |
Db 2 TYKVYPWG 9

RESULT 6
US-09-207-844-3
Sequence 3, Application US/09207844
Patent No. 5968774
GENERAL INFORMATION:
APPLICANT: Lehmbeck, Jan
TITLE OF INVENTION: Host Cell Expressing Reduced Levels
TITLE OF INVENTION: Of A Metalloprotease And Methods Using The Host Cell In Protei
TITLE OF INVENTION: Production
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5968774o No. 5968774disk of No. 5968774th America, Inc.
STREET: 405 Lexington Avenue, Suite 6400
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/207,844
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/894,772
FILING DATE: 27-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Gregg, Valeta A.
REGISTRATION NUMBER: 35,127
REFERENCE/DOCKET NUMBER: 4300.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-867-0298
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-207-844-3

Query Match 34.9%; Score 29; DB 2; Length 14;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 TYKVYMSG 15
| | | | |
Db 2 TYKVYPWG 9

RESULT 7
PCT-US95-05534-2
Sequence 2, Application PC/TUS9505534
GENERAL INFORMATION:
APPLICANT:
APPLICANT:

;; TITLE OF INVENTION: NOVEL METALLOPROTEASE HAVING INCREASED
;; TITLE OF INVENTION: ACTIVITY
;; NUMBER OF SEQUENCES: 10
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Novo Nordisk of North America, Inc.
;; STREET: 405 Lexington Avenue, 64th Floor
;; CITY: New York
;; STATE: New York
;; COUNTRY: USA
;; ZIP: 10174-6401
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: tape
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Ver. 1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: PCT/US95/05534
;; FILING DATE:
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; FILING DATE: 04-MAY-1994
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; FILING DATE: 03-MARCH-1995
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Agnis Dr., Cheryl H.
;; REGISTRATION NUMBER: 34,086
;; REFERENCE/DOCKET NUMBER: 4180.204-WO
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 212-867-0123
;; TELEFAX: 212-867-9655
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 14 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
PCT-US95-05534-2

Query Match 34.9%; Score 29; DB 5; Length 14;
Best Local Similarity 75.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 TYKVYMSG 15
DB 2 TYKVYPWG 9

RESULT 8
US-09-528-200-40
; Sequence 40, Application US/09528200
; Patent No. 6630570
; GENERAL INFORMATION:
; APPLICANT: LICHA, KAI
; APPLICANT: BECKER, ANDREAS
; APPLICANT: SEMMLER, WOLFGANG
; APPLICANT: WEIDENMANN, BERTRAM
; APPLICANT: HESSNIUS, CARSTEN
; APPLICANT: VOLKMER-ENGERT, RUDOLF
; APPLICANT: SCHNEIDER-MERGENER, JENS
; APPLICANT: BHARGAVA, SARAH
; TITLE OF INVENTION: SHORT-CHAIN PEPTIDE-DYE CONJUGATES AS CONTRAST MEDIA
; FILE REFERENCE: SCH-1731
; CURRENT APPLICATION NUMBER: US/09/528.200
; CURRENT FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: DE 199 17 713.9
; PRIOR FILING DATE: 1999-09-04
; NUMBER OF SEQ ID NOS: 196

;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 40
;; LENGTH: 12
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-528-200-40

Query Match 33.7%; Score 28; DB 4; Length 12;
Best Local Similarity 41.7%; Pred. No. 54;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 IAKQWYKVMYS 14
DB 1 LRQWAVKKYL 12

RESULT 9
US-08-413-708B-6
; Sequence 6, Application US/08413708B
; Patent No. 5972883
; GENERAL INFORMATION:
; APPLICANT: GOZES, ILLANA
; APPLICANT: FRIDKIN, Matityahu
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING
; TITLE OF INVENTION: VIP, OR ANALOGUES DERIVATIVES AND FRAGMENTS THEREOF
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,708B
; FILING DATE: 30-MAR-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/207,671
; FILING DATE: 09-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: IL 105061
; FILING DATE: 16-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWDY, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: GOZES=3A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-413-708B-6

Query Match 33.7%; Score 28; DB 2; Length 13;
Best Local Similarity 38.5%; Pred. No. 59;
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 5 KQMTYKVMYSGTV 17
DB 1 KQMAVKKYLAAYL 13

```
RESULT 10
US-09-528-200-33
; Sequence 33, Application US/09528200
; Patent No. 6630570
; GENERAL INFORMATION:
; APPLICANT: LICHIA, KAI
; APPLICANT: BECKER, ANDREAS
; APPLICANT: SEMMLER, WOLFHARD
; APPLICANT: WEIDENMANN, BERTRAM
; APPLICANT: HESSNIUS, CARTSEN
; APPLICANT: VOLKMER-ENGERT, RUDOLF
; APPLICANT: SCHNEIDER-MERGNER, JENS
; APPLICANT: BHARGAVA, SARAH
; TITLE OF INVENTION: SHORT-CHAIN PEPTIDE-DYE CONJUGATES AS CONTRAST MEDIA
; TITLE OF INVENTION: FOR OPTICAL DIAGNOSIS
; FILE REFERENCE: SCH-1731
; CURRENT APPLICATION NUMBER: US/09/528,200
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: DE 199 17 713.9
; PRIOR FILING DATE: 1999-09-04
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 33
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-528-200-33

Query Match      33.7%; Score 28; DB 4; Length 13;
Best Local Similarity 41.7%; Pred. No. 59;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      3 IAKQMTYKYVMS 14
      : ||| | |::
Db      1 LRQMVAVKYLN 12

RESULT 11
US-09-528-200-39
; Sequence 39, Application US/09528200
; Patent No. 6630570
; GENERAL INFORMATION:
; APPLICANT: LICHIA, KAI
; APPLICANT: BECKER, ANDREAS
; APPLICANT: SEMMLER, WOLFHARD
; APPLICANT: WEIDENMANN, BERTRAM
; APPLICANT: HESSNIUS, CARTSEN
; APPLICANT: VOLKMER-ENGERT, RUDOLF
; APPLICANT: SCHNEIDER-MERGNER, JENS
; APPLICANT: BHARGAVA, SARAH
; TITLE OF INVENTION: SHORT-CHAIN PEPTIDE-DYE CONJUGATES AS CONTRAST MEDIA
; TITLE OF INVENTION: FOR OPTICAL DIAGNOSIS
; FILE REFERENCE: SCH-1731
; CURRENT APPLICATION NUMBER: US/09/528,200
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: DE 199 17 713.9
; PRIOR FILING DATE: 1999-09-04
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-528-200-39

Query Match      33.7%; Score 28; DB 4; Length 13;
Best Local Similarity 41.7%; Pred. No. 59;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      3 IAKQMTYKYVMS 14
      : ||| | |::
Db      2 LRQMVAVKYLN 13

RESULT 12
US-09-528-200-26
; Sequence 26, Application US/09528200
; Patent No. 6630570
; GENERAL INFORMATION:
; APPLICANT: LICHIA, KAI
; APPLICANT: BECKER, ANDREAS
; APPLICANT: SEMMLER, WOLFHARD
; APPLICANT: WEIDENMANN, BERTRAM
; APPLICANT: HESSNIUS, CARTSEN
; APPLICANT: VOLKMER-ENGERT, RUDOLF
; APPLICANT: SCHNEIDER-MERGNER, JENS
; APPLICANT: BHARGAVA, SARAH
; TITLE OF INVENTION: SHORT-CHAIN PEPTIDE-DYE CONJUGATES AS CONTRAST MEDIA
; TITLE OF INVENTION: FOR OPTICAL DIAGNOSIS
; FILE REFERENCE: SCH-1731
; CURRENT APPLICATION NUMBER: US/09/528,200
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: DE 199 17 713.9
; PRIOR FILING DATE: 1999-09-04
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-528-200-26

Query Match      33.7%; Score 28; DB 4; Length 14;
Best Local Similarity 41.7%; Pred. No. 65;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      3 IAKQMTYKYVMS 14
      : ||| | |::
Db      1 LRQMVAVKYLN 12

RESULT 13
US-09-528-200-32
; Sequence 32, Application US/09528200
; Patent No. 6630570
; GENERAL INFORMATION:
; APPLICANT: LICHIA, KAI
; APPLICANT: BECKER, ANDREAS
; APPLICANT: SEMMLER, WOLFHARD
; APPLICANT: WEIDENMANN, BERTRAM
; APPLICANT: HESSNIUS, CARTSEN
; APPLICANT: VOLKMER-ENGERT, RUDOLF
; APPLICANT: SCHNEIDER-MERGNER, JENS
; APPLICANT: BHARGAVA, SARAH
; TITLE OF INVENTION: SHORT-CHAIN PEPTIDE-DYE CONJUGATES AS CONTRAST MEDIA
; TITLE OF INVENTION: FOR OPTICAL DIAGNOSIS
; FILE REFERENCE: SCH-1731
; CURRENT APPLICATION NUMBER: US/09/528,200
; CURRENT FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: DE 199 17 713.9
; PRIOR FILING DATE: 1999-09-04
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-528-200-32

Query Match      33.7%; Score 28; DB 4; Length 14;
Best Local Similarity 41.7%; Pred. No. 65;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      3 IAKQMTYKYVMS 14
      : ||| | |::
Db      1 LRQMVAVKYLN 12
```

QY 3 IAKQMTYKVYMS 14
: ||| |::
Db 2 LRQMAVKKYL 13

RESULT 14

US-08-486-721A-16
; Sequence 16, Application US/08486721A
; Patent No. 5739025
; GENERAL INFORMATION:
; APPLICANT: Fukazawa, Chikafusa
; TITLE OF INVENTION: Method of Producing
; TITLE OF INVENTION: Asparaginyl Endoprotease
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frischauf, Holtz, Goodman, Langer & Chick
; STREET: 767 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10017-2023
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3+ inch, 1.4 mb
; COMPUTER: IBM PC
; OPERATING SYSTEM: MS DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,721A
; FILING DATE: 7-JUNE-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/091,991
; FILING DATE: 12-JULY-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Barth, Richard S.
; REGISTRATION NUMBER: 28,180
; REFERENCE/DOCKET NUMBER: 930587/HG
; TELEPHONE: (212) 319-4900
; TELEFAX: (212) 319-5101
; TELEX: 236268
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE:
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM:
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:

; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
US-08-486-721A-16

Query Match 33.7%; Score 28; DB 1; Length 15;
Best Local Similarity 41.7%; Pred. No. 70;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 IAKQMTYKVYMS 14
: ||| |::
Db 4 LRQMAVKKYL 15

RESULT 15

US-09-528-200-19
; Sequence 19, Application US/09528200
; Patent No. 6630570
; GENERAL INFORMATION:
; APPLICANT: LICHA, KAI
; APPLICANT: BECKER, ANDREAS
; APPLICANT: SEMMLER, WOLFHARD
; APPLICANT: WEIDENMANN, BERTRAM
; APPLICANT: HESSNIUS, CARTSEN
; APPLICANT: VOLKMER-ENGERT, RUDOLF
; APPLICANT: SCHNEIDER-MERGENER, JENS
; APPLICANT: BHARGAVA, SARAH
; TITLE OF INVENTION: SHORT-CHAIN PEPTIDE-DYE CONJUGATES AS CONTRAST MEDIA
; FILE REFERENCE: SCH-1731
; CURRENT APPLICATION NUMBER: US/09/528,200
; CURRENT FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: DE 199 17 713.9
; PRIOR FILING DATE: 1999-09-04
; NUMBER OF SEQ ID NOS: 196
; SEQ ID NO 19
; SOFTWARE: PatentIn Ver. 2.1
; LENGTH: 15
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-528-200-19

Query Match 33.7%; Score 28; DB 4; Length 15;
Best Local Similarity 41.7%; Pred. No. 70;
Matches 5; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 3 IAKQMTYKVYMS 14
: ||| |::
Db 1 LRQMAVKKYL 12

Search completed: August 12, 2004, 07:06:30
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 06:12:47 ; Search time 21.3012 Seconds
(without alignments)
1043.144 Million cell updates/sec

Title: US-09-890-463-3

Perfect score: 1268

Sequence: 1 SVIAKQWTKYVMSTGVNGH.....SIARKPLVACCFRVRKSRHK 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	184.5	14.6	238	1 JQ1514	green-fluorescent
2	106	8.4	583	2 S64909	probable membrane
3	90	7.1	752	2 S23818	hypothetical prote
4	87	6.9	787	2 E82323	organic solvent to
5	86.5	6.8	725	1 IJWSNG	neural cell adhesi
6	86	6.8	770	2 S00643	anthranilate synth
7	85	6.7	260	2 B70419	hypothetical prote
8	84	6.6	1162	2 S14939	E2 glycoprotein pr
9	83.5	6.6	390	2 D64903	probable arylsulfa
10	83.5	6.6	3461	2 S58870	reelin precursor -
11	82.5	6.5	876	2 A89944	alanyl-tRNA synth
12	82	6.5	248	2 S58096	hypothetical prote
13	82	6.5	1162	2 S14940	E2 glycoprotein pr
14	81.5	6.4	390	2 F90891	probable enzyme [1
15	81.5	6.4	390	2 C85726	probable enzyme [2
16	80	6.3	1822	2 S63982	collagen alpha 2 c
17	79	6.2	334	2 D95982	hypothetical expor
18	79	6.2	403	2 T51828	probable photosyst
19	79	6.2	687	1 B32382	ubiquinol-cytochro
20	78.5	6.2	363	2 S30149	cysteine proteinase
21	78.5	6.2	1254	2 S46636	hypothetical prote
22	77.5	6.1	266	1 CDFJ13	chlorophyll a/b-bi
23	77.5	6.1	269	1 G65102	probable transcript
24	77.5	6.1	269	2 A91130	robable transcript
25	77.5	6.1	269	2 A85975	hypothetical prote
26	77.5	6.1	333	2 B64380	conserved hypothet
27	77.5	6.1	513	2 D69824	lipoxigenase AtLOX
28	77.5	6.1	870	2 T47454	lipoxigenase (EC 1
29	77.5	6.1	896	2 JQ2391	

ALIGNMENTS

RESULT 1

JQ1514

green-fluorescent protein [validated] - hydromedusa (Aequorea victoria)

C:Species: Aequorea victoria

C>Date: 03-Dec-1999 #sequence revision 03-Dec-1999 #text change 23-Mar-2001

C:Accession: J50692; JQ1514; FQ0335; S48693; S51330; S51331

R:Prasher, D.C.; Eckenrode, V.K.; Ward, W.W.; Prendergast, F.G.; Cormier, M.J.

Gene 111, 229-233, 1992

A>Title: Primary structure of the Aequorea victoria green-fluorescent protein.

A:Reference number: JQ1514; MUID:92175527; PMID:1347277

A:Accession: J50692

A:Molecule type: DNA

A:Residues: 1-107, 'S', 109-238 <PRA1>

A:Cross-references: GB:M62654; NID:gl55662; PIDN:AAA27722.1; PID:gl55663

A:Accession: JQ1514

A:Molecule type: mRNA

A:Residues: 1-99, 'P', 101-140, 'L', 142-218, 'V', 220-238 <PRA2>

A:Cross-references: GB:M62653; NID:gl55660; PIDN:AAA27721.1; PID:gl55661

A:Accession: FQ0335

A:Molecule type: protein

A:Residues: 46-64; 74-122; 132-151; 154-183; 185-200 <PRA3>

R:Inouye, S.; Tsuji, F.I.

FEBS Lett. 351, 211-214, 1994

A>Title: Evidence for redox forms of the Aequorea green fluorescent protein.

A:Reference number: S48693; MUID:94364470; PMID:8082767

A:Accession: S48693

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-24, 'Q', 26-156, 'P', 158-171, 'K', 173-238 <INO>

A:Cross-references: GB:I29345; NID:g606383; PIDN:AAA58246.1; PID:g606384

R:Watkins, J.N.; Campbell, A.K.

submitted to the EMBL Data Library, January 1995

A:Reference number: S51330

A:Accession: S51330

A:Molecule type: mRNA

A:Residues: 1-13, 'V', 15-24, 'Q', 26-44, 'N', 46-153, 'G', 155-156, 'P', 158-171, 'K', 173-227, 'R', 2

A:Cross-references: EMBL:X83959; NID:g634008; PIDN:CAA58789.1; PID:g634009

A:Experimental source: clone gfp1

A:Accession: S51331

A:Molecule type: mRNA

A:Residues: 1-24, 'Q', 26-29, 'R', 31-83, 'L', 85-153, 'G', 155-156, 'P', 158-171, 'K', 173-208, 'Q', 2

A:Cross-references: EMBL:X83960; NID:g634010; PIDN:CAA58790.1; PID:g634011

A:Experimental source: clone gfp2

R:Yang, F.; Moss, L.G.; Phillips Jr., G.N.

submitted to the Brookhaven Protein Data Bank, August 1996

A:Reference number: A65692; PDB:1GLF

A:Contents: annotation; X-ray crystallography, 1.9 angstroms, residues 'A', 2-79, 'R', 81-95

A>Note: engineered sequence based on JQ1514, cloned and expressed in Escherichia coli

R:Yang, F.; Moss, L.G.; Phillips Jr., G.N.

Nat. Biotechnol. 14, 1246-1251, 1996

A>Title: The molecular structure of green fluorescent protein.

A:Reference number: A58953; MUID:98294543; PMID:9631087

A;Contents: annotation; X-ray crystallography, 1.9 angstroms
C;Comment: This protein is excited by the photoprotein aequorin (see PIR:AQJFNV) emitting
C;Comment: The chromophore of this protein is formed by modification of Ser-dehydro-Tyr-
C;Genetics:
A;Gene: GFP

A;Introns: 69/3; 167/3

C;Superfamily: green-fluorescent protein

C;Keywords: chromoprotein; luminescence

F;65-67/Cross-link: 5-imidazolinone (Ser-Gly) #status experimental

F;66/Modified site: dehydrotyrosine (Tyr) #status experimental

Query Match 14.6%; Score 184.5; DB 1; Length 238;
Best Local Similarity 24.8%; Pred. No. 6.4e-09;
Matches 51; Conservative 47; Mismatches 91; Indels 17; Gaps 8;

QY 11 VMSTGVNGHVFVEGDKGKPYEGEQTVRLAVTKGGPLPFADWILSPQCQVGSIPFTKY 70
DB 16 VELDGVNKHFSVSGEGEGATYKGLTKICTT-GKLPVWPPLVTFTSVGVQCFGRY 74
QY 71 PEDIP-DYVKQSPFGRYTWERIMNFEDGAVCTVSDSSIQNCFIYHVKFSGLNFPNG 128
DB 75 PDHMKQHDFFKSAPEGVGVQERTIFYKDDGNVKTAEVKFEQDTLVNRIELKIDFKEDG 134
QY 129 PVMQKKTQGWPNTERLP-----ARDGMLIGNFP-MALKLEGGGHYLCEP--KSTYKARK 180
DB 135 NILGHKME-YNYSNHNVMADKQNGIKV--NFKIRHNIEDGVSQVLADHYQQNTPIGDG 191
QY 181 PVKMPGVHYVDRKLDVT---NHNKDY 203
DB 192 PVLIPDNHYLTQSALSKDPNEKRDH 217

RESULT 2

S64909
Probable membrane protein YLR077w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein L2349
C;Species: Saccharomyces cerevisiae
C;Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 19-Apr-2002
C;Accession: S64909
R;Pohl, T.M.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64909
A;Accession: S64909
A;Molecule type: DNA
A;Residues: 1-583 <POH>
A;Cross-references: EMBL:Z73249; NID:gl360436; PID:e245794; PID:gl360437; GSPDB:GNO0012;
A;Experimental source: strain S288C
C;Genetics:
A;Gene: MIPS:YLR077w
A;Cross-references: SGD:S0004067
A;Map position: 12R
C;Superfamily: Saccharomyces cerevisiae
C;Keywords: transmembrane protein
F;81-97/Domain: transmembrane #status predicted <TMM>

Query Match 8.4%; Score 106; DB 2; Length 583;
Best Local Similarity 21.0%; Pred. No. 0.18;
Matches 56; Conservative 33; Mismatches 88; Indels 90; Gaps 13;

QY 18 NGHYFEVEGDKGKPYEGEQTVR-----LAVTKGGPL-PFADWILSPQCQY 62
DB 349 NNQLFEILLNFKF-HEGEDVVRKREIKKIACGSVHTLAIDTKTGSIYAFGWN-----RF 401
QY 63 G-----SIFFTKYPEDIDYVKQSPFGRYTWERI----- 91
DB 402 QQLALPISYNLEYVSPFRSVTHAFKPPFGMTNKKCVDIHCDDETSFVYIRKPGSTSDH 461
QY 92 -MNFEDGAVCTVSDSSIQNCFIYHVKFSGLNFPNGFVMOKKTQGWEPNTERLFA--- 147
DB 462 YFAFGNGLFGLGNNTFNKNSQCDPIKIKSD-----DKKLTNMSCGSHCVFTETE 510
QY 148 --RDGMLTGNFMALKLEGGH-YLCEFKSTYKARKPKVM-----PGVHYVDRKLDVTNH 199

DB 511 QENEVIAGNN-----DHGQLGIGKTKMKCAKPMNIPVLKFGQDTTD--LDSIYN 559
QY 200 NKDYTSVEQREISIAARKPLVACCPFRV 226
DB 560 SKLHLKKEQQRVVTNGNK---SCLYWRV 583

RESULT 3

S23818
hypothetical protein Tnp2 - garden snapdragon transposable element Tam1
C;Species: Antirrhinum majus (garden snapdragon)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 13-Sep-1998
C;Accession: S23818; S16551
R;Nacken, W.K.F.; Piotrowiak, R.; Saedler, H.; Sommer, H.
submitted to the EMBL Data Library, January 1991
A;Description: The transposable element Tam1 from Antirrhinum majus shows structural homology
A;Reference number: S23817
A;Accession: S23818
A;Molecule type: DNA
A;Residues: 1-752 <NAC>
A;Cross-references: EMBL:X57297
R;Nacken, W.K.F.; Piotrowiak, R.; Saedler, H.; Sommer, H.
Mol. Gen. Genet. 228, 201-208, 1991
A;Title: The transposable element Tam1 from Antirrhinum majus shows structural homology
A;Reference number: S16551; MUID:91360065; PMID:1715971
A;Accession: S16551
A;Molecule type: DNA
A;Residues: 29-752 <NAC>
A;Cross-references: EMBL:X57297
C;Genetics:
A;Mobile element: transposable element Tam1
A;Start codon: GTT

Query Match 7.1%; Score 90; DB 2; Length 752;
Best Local Similarity 23.2%; Pred. No. 6.3;
Matches 53; Conservative 40; Mismatches 79; Indels 56; Gaps 14;

QY 31 KPYEGEQTVRLAVTKGGPLPFADWILSPQCQVGSIPFTKYPEDIDYVKQSPFGRYTWER 90
DB 483 KSFDGNEYRRA-----PIALTGDMVSEITGFNIKFGKVVDDNF-----TLP--LNWKK 530
QY 91 IMNFEDGAVCTVSDSSIQNCFIYHVKFSGLNFPNP-----GPMOKKTQGWEPNTER 144
DB 531 RSIFPD---LPYWKDSLRLHNFVWHIE-----KNVCESIIIGTLNLEGRTKDHENS 580
QY 145 LFARDGMLIGNFMALKLEGGGHYL---C-----EFKSTYKARKPKVMP-GY-HYVDRK 193
DB 581 LDLKD-MGIRSELHPISLESGRHYLPAACYSMSKKEKEIVFEILKTVKVPDGIASNISRR 639
QY 194 LDVTNH-----NKDYTSVEQREISIA-----RKPLV-ACCFER 225
DB 640 VOLKPNKISGLKSHDHHILMQLLPIALRKVLPKHVRTPLIKLCTFFR 687

RESULT 4

E82323
organic solvent tolerance protein VC0446 [imported] - Vibrio cholerae (strain N16961 sero-
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: E82323
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: E82323
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-787 <HEI>
A;Cross-references: GB:AE004131; GB:AE003852; NID:g9654856; PIDN:AAF93619.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:

A;Gene: VC0446

A;Map position: 1

C;Superfamily: organic solvent tolerance protein

Query Match 6.9%; Score 87; DB 2; Length 787;
Best Local Similarity 19.2%; Pred. No. 12;
Matches 45; Conservative 33; Mismatches 92; Indels 64; Gaps 9;

QY 6 QMTYKVMSTGVN-----GHYEVEVDGKGKPYEGEQTVRLAVTKGGLPFA--WDIL 56

Db QLSNYYAPETMKYLDLVLVSHVSRFETDARGK-----SATRVHIEPLKTKPFNTWGNW 443

QY 57 SPQCYGSGIPTKYPEDIPYVKQSPGRYTWERIMNPFEDGAVCTVSNDSIQGNCFIYH 116

Db 444 TTEAR---VLGTYYQDLDKTID-----AKLEESVTRVPIRSV----- 480

QY 117 VKFGLNPPNPVQKQTKQGWENTERLFA-----RDGMLIGNPMAL----- 160

Db 481 ---AGIVLERDTVLDDYTQLEPKIQVLYVPEKYQDNIGLYDSTLLQTDYVGLFRSKY 537

QY 161 ---KLEGG-----CHYLCEPKSTYKARKPKVMPGYHYVDRKLDVTHNKKDYS 205

Db 538 SGVDRIESANQVSYGASTRFEDSNYKERLNTAFQGIYLDLSKIMPSNKNPDSTS 591

RESULT 5

LUMSNG

neural cell adhesion molecule 1 precursor, GPI-anchored splice form - mouse

N;Alternate names: NCAM-120

C;Species: Mus musculus (house mouse)

C;Date: 31-Mar-1993 #sequence revision 31-Mar-1993 #text_change 31-Dec-2000

C;Accession: A29673; S00382; A44290

R;Barthels, D.; Santoni, M.J.; Wille, W.; Ruppert, C.; Chaix, J.C.; Hirsch, M.R.; Pontec

EMBO J. 6, 907-914, 1987

A;Title: Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000

A;Reference number: A29673; MUID:87246524; PMID:3595563

A;Accession: A29673

A;Molecule type: mRNA

A;Residues: 1-725 <BAR>

A;Cross-references: EMBL:X00051; NID:953342; PIDN:CAA68263.1; PID:953343

R;Barbas, J.A.; Chaix, J.C.; Steinmetz, M.; Goridis, C.

EMBO J. 7, 625-632, 1988

A;Title: Differential splicing and alternative polyadenylation generates distinct NCAM t

A;Reference number: S00382; MUID:88283628; PMID:3396534

A;Accession: S00382

A;Molecule type: DNA

A;Residues: 642-656, D', 658-725 <BA2>

A;Cross-references: EMBL:X07195

R;Rougon, G.; Marshak, D.R.

J. Biol. Chem. 261, 3396-3401, 1986

A;Title: Structural and immunological characterization of the amino-terminal domain of n

A;Reference number: A44290; MUID:86140120; PMID:3512556

A;Accession: A44290

A;Molecule type: protein

A;Residues: 20-36 <ROU>

C;Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol

C;Comment: Several forms of NCAM are produced by alternative splicing. See also PIR:LU

C;Genetics:

A;Gene: NCAM

A;Map position: 9

A;Introns: 701/1

C;Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu

F;1-15/Domain: signal sequence #status predicted <SIG>

F;34-98/Domain: immunoglobulin homology <IMM1>

F;132-191/Domain: immunoglobulin homology <IMM2>

F;152-156/Region: heparin binding #status predicted

F;161-165/Region: heparin binding #status predicted

F;228-290/Domain: immunoglobulin homology <IMM3>

F;263-272/Region: NCAM binding #status predicted

F;323-388/Domain: immunoglobulin homology <IMM4>

F;420-482/Domain: immunoglobulin homology <IMM5>

F;519-596/Domain: fibronectin type III repeat homology <FN3A>

F:625-685/Domain: fibronectin type III repeat homology <FN3B>

F:41-96,139-189,235-288,330-386,427-480/Disulfide bonds: #status predicted

F:222,316,348,424,450,479/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.8%; Score 86.5; DB 1; Length 725;
Best Local Similarity 21.7%; Pred. No. 12;
Matches 53; Conservative 41; Mismatches 89; Indels 61; Gaps 15;

QY 10 KVMSTGVNHYFVEVDGKGK--PYE-----GEQTVR-LAVTKGGLPFAWDILSP- 58

Db 455 KIY--NTPSASYLEVTPDSENFQNYNCTAVNRIGQESLEFILVQADTPSPSIDRVEPY 512

QY 59 ----QCOY-----GSTPFTKYPEDIPDYVKQSPGRYTWERIMNPFEDGAVCTVSNDSI 108

Db 513 SSTAQVQFDEPTEATGGVPILKYKAWSLGEESW--HFTWYD-----AKEANM 558

QY 109 QGNCFI-----YHVKFGSLNPPNPVQVQ---KKTQG---WEPTERLFAWDGLMI 153

Db 559 EGIVTIMGKPEPTYSDRLAALNGKLGELIQPSESCTQVPPELSAPKLEQMGEDGNSI 618

QY 154 GNNFMALKLEGG-----HYLCEPKSTYKARKP-VKMP--GYHYVDRKLDVTHNKKDYS 206

Db 619 KVNIL--TKQDDGGSPIRHYLVKYRALASEWKPEIRLPSGSHVLMKSLDWAIEYVYVA 676

QY 207 EORE 210

Db 677 ENQQ 680

RESULT 6

S00643

anthranilate synthase multifunctional protein - Aspergillus niger

N;Alternate names: gene trpG-trpC-trpF protein; glutamine amidotransferase

N;Contains: anthranilate synthase (EC 4.1.3.27); indole-3-glycerol-phosphate synthase (EC

C;Species: Aspergillus niger

C;Date: 07-Sep-1990 #sequence revision 07-Sep-1990 #text_change 31-Mar-2000

C;Accession: S00643; A23979; B23979

R;Kos, T.; Kuijvenhoven, A.; Hessing, H.G.M.; Pouwels, P.H.; van den Hondel, C.A.M.J.J.

Curr. Genet. 13, 137-144, 1988

A;Title: Nucleotide sequence of the Aspergillus niger trpC gene: structural relationship

A;Reference number: S00643; MUID:88223483; PMID:2836085

A;Accession: S00643

A;Molecule type: DNA

A;Residues: 1-770 <KOS>

A;Cross-references: EMBL:X07071; NID:92420; PIDN:CAA30107.1; PID:92421

R;Kos, A.; Kuijvenhoven, J.; Wernars, K.; Bos, C.J.; van den Broek, H.W.J.; Pouwels, P.H.

Gene 39, 231-238, 1985

A;Reference number: A91539; MUID:86137391; PMID:2936650

A;Accession: A23979

A;Molecule type: DNA

A;Residues: 1-69 <KO2>

A;Cross-references: GB:M14404; NID:9166536; PIDN:AAA32709.1; PID:9166537

A;Accession: B23979

A;Molecule type: DNA

A;Residues: 392-433 <KO3>

A;Cross-references: GB:M14403; NID:9166538; PIDN:AAA32710.1; PID:9166539

C;Genetics:

A;Gene: trpC

C;Superfamily: trpG-trpC-trpF trifunctional enzyme; trpC homology; trpF homology; trpG hc

C;Keywords: carbon-carbon lyase; carboxy-lyase; intramolecular lyase; isomerase; oxo-acid

F;25-219/Domain: glutamine amidotransferase #status predicted <GAT>

F;26-216/Domain: trpC homology <TRG>

F;255-518/Domain: trpC homology <TRC>

F;255-514/Domain: indole-3-glycerol-phosphate synthase #status predicted <IGPS>

F;537-767/Domain: trpF homology <TRF>

F;545-770/Domain: N-(5'-phosphoribosyl)anthranilate isomerase #status predicted <PRAI>

F;104/Active site: Cys #status predicted

Query Match 6.8%; Score 86; DB 2; Length 770;

Best Local Similarity 24.3%; Pred. No. 15;

Matches 50; Conservative 23; Mismatches 61; Indels 72; Gaps 12;

QY 10 KVMSTGV-NGHYFEVDGKGKPYEGQTVRLAVTKGGLPFAWDILSPQCYGSIPT 68

reelin precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 21-Jul-2003
A/Accession: S58870; S71844; I49297
R/D/Arcangelo, G.; Mao, G.G.; Chen, S.C.; Soares, H.D.; Morgan, J.I.; Curran, T.
Nature 374, 719-723, 1995
A/Title: A protein related to extracellular matrix proteins deleted in the mouse mutant
A/Reference number: I49297; MUID:95231649; PMID:7715725
A/Accession: S58870
A/Status: nucleic acid sequence not shown
A/Molecule type: mRNA
A/Residues: 1-3461 <DAR>
A/Cross-references: EMBL:U24703; NID:G902486; PID:G902487
R/D/Arcangelo, G.
submitted to the EMBL Data Library, April 1995
A/Reference number: S71844
A/Accession: S71844
A/Molecule type: mRNA
A/Residues: 1-215, T, 217-1905, 'S', 1907-3355, 'V', 3357-3391, 'N', 3393-3461 <DA2>
A/Cross-references: EMBL:U24703; NID:G902486; PID:G902487
F/I-27/Domain: signal sequence #status predicted <SIG>
F/28-3461/Product: reelin #status predicted <MAT>
F/1769-1795/Domain: EGF homology <EGF>

Query Match	6.6%; Score 83.5; DB 2; Length 3461;
Best Local Similarity	20.3%; Pred. No. 1.6e+02;
Matches	57; Conservative 24; Mismatches 75; Indels 125; Gaps 12;
Qy	12 YMSGTVNGHYFEVEGDGK---GKPYEGEQTVRLAVTKG---GPLFPANDI-----LSQP 59
Db	2080 YYACTGTGWRREVYHFKLHLCG-----SVRFWYGGFYPAQSQVTVWADNVITGPQ 2132
Qy	60 CQ---YG-----SIPFTKYPEDIPDYVKQSFPGRYTWERIMNF-----94
Db	2133 CEEMCYGHGSCINGTKICIDPGYGGPTCKISTKNPDLKDDFBGQLSDRFLMLSGGKPS 2192
Qy	95 -----EDGACVTGNSDSSIOGNCFIYHVKF-----SGLNFPNPGP 129
Db	2193 RKCGLSSGNMLFFNEDGLRMLVTRDLDS-----HARFVQFPMRLCCGKGVDPDRSQP 2246
Qy	130 VM-----OQKKTQGWEPNTERLFP-----146
Db	2247 VLLQVSLNGGLSWSLLQELFLSNSNVGRYIALEMLKARSGSTRLWQVPSENGHYFSP 2306
Qy	147 -ARGCMILIGNFMALKUEGGGHYICEFKSTYKARKPVMPG 186
Db	2307 WVIDQILIGNNI-----SGNTVLDDGFSITLDSRKWILLHPG 2341

RESULT 11

A89944
alanyl-tRNA synthetase [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 17-Mar-2003
C:Accession: A89944
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: A89944
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-876 <R>
A:Cross-references: GB:BA000018; PID:g13701416; PIDN:BAE42710.1; GSPDB:GNC00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: alas
C:Superfamily: alanyl-tRNA ligase

Query Match 6.5%; Score 82.5; DB 2; Length 876;
Best Local Similarity 21.9%; Pred. No. 35;

Matches	46;	Conservative	25;	Mismatches	60;	Indels	79;	Gaps	11;
Qy	16	TVNGH-YFEVEGDGKGRPYEGEQTURLAVTKGGLPFFAWDILSPQCQGSIP----	FTKY	70					
Db	81	TARHHTTFEMLGNFSIGDYFQEALE-----	FAWEFL	130					
Qy	71	PEDIPDYVKQSPGRYTHWRIMTFEDGAVCTVSDSSICQNCFIYHVKFSGLNFPENGPV	130						
Db	131	PEDMEAY-----	NWKKDKIGLESRII-----	RIEGN-----	FWDIGEGPSG--	167			
Qy	131	MQKKTQGWEPNTERLFFARDGMLIGNNFWALKLEGGGHYLCFPGSTYKARKPKVMFGYHV	190						
Db	168	-----	PNTEIFYDR-GEAYGQDDPAEEMYPGEN-----	-----	-----	195			
Qy	191	DRKLDV-----	TNNHKD--YTSVEQREI	211					
Db	196	ERYLEVWNLVFESEFNHNKDKHSYTLPLNENI	225						

RESULT 12

S58096
hypothetical protein SPAC13C5.04 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 21-Jan-2000
C;Accession: S58096; T37617
R;Devlin, K.; Churcher, C.M.
submitted to the EMBL Data Library, July 1995
A;Reference number: S58093
A;Accession: S58096
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-248 <DEV>
A;Cross-references: EMBL:Z50112; NID:g908898; PIDN:CAA90455.1; PID:g908893
R;Devlin, K.; Churcher, C.M.; Barrall, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, July 1995
A;Reference number: Z21731
A;Accession: T37617
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-248 <DE2>
A;Cross-references: EMBL:Z50112; PIDN:CAA90455.1; GSPDB:GN00066; SPDB:SPAC13C5.04
A;Experimental source: strain 972h-; cosmid cl3C5
C;Genetics:
A;Gene: SPAC13C5.04
A;Map position: 1

Query Match	6.5%;	Score 82;	DB 2;	Length 248;
Best Local Similarity	24.2%;	Fred. NO. 7.9;		
Matches 45;	Conservative 29;	Mismatches 62;	Indels 50;	Gaps 12;
QY	69	KYPEDIPYVWQSPFGRYTWERIMN--FEDGAVCTVSND-----SSIQNCFTY-HV	117	
Db	54	KNPNDDYQ--KEDFPN-----INAILITGSKASATSDAPWIKKLSFVKDVLKPYPHI	104	
QY	118	KFSGLNF-----PPNGPVWQKKTQGWHE-----PNTERLFARDGMLIG--NNFMA	159	
Db	105	KIVGLCFGHQIVAKAAGVPITIQ-NPKGWESVTVVQLTENGEKFFGKRVKININQMHQDMA	163	
QY	160	LKLEGGGHYL-----CEPKSYKARKPKVMPGY-----HYVDRKLDVTNNHKDYTSVEOR	209	
Db	164	VDVPEGFELLGSTECECFQIYKPKQALTFQGHPEFSTEVVNTWVKVLRGTEVFTE-QQK	222	
QY	210	EISLAR	215	
Db	223	BEALKR	228	

RESULT 13
S14940
E2 glycoprotein precursor - avian infectious bronchitis virus (strain M42-S)
N:Alternate names: peplomer glycoprotein
C:Species: avian infectious bronchitis virus, IBV
A:Variety: strain M42-S

C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C;Accession: S14940
R;Niesters, H.G.M.; Lenstra, J.A.; Spaan, W.J.M.; Zijderveld, A.J.; Bleumink-Pluym, N.M.
Virus Res. 5, 253-263, 1986
A;Title: The peplomer protein sequence of the M41 strain of coronavirus IBV and its comp
A;Reference number: S07421; MUID:87021475; PMID:2429473
A;Accession: S14940
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: genomic RNA
A;Residues: 1-1162 <NIE>
C;Superfamily: coronavirus E2 glycoprotein

Query Match 6.4%; Score 82; DB 2; Length 1162;
Best Local Similarity 20.7%; Pred. No. 55;
Matches 60; Conservative 29; Mismatches 65; Indels 136; Gaps 16;

Qy 3 IAKQWTK-----VMSG-----TNGHYFEVEGDGKGPYGEQTVR 40
Db VAKYPTSFQCVNLTSTVYNGDLVYTSNFTDVTSGAVYFKA----- 192

Qy 41 LAVTKGGPLP-----AWDIL-----SPQ-----COYGSIPFTK--YPRDI 74
Db -----GGPIYKWKREVKALAYFVNGTAQDVILCDGSPRGLLACQYNTGNFSDGFPPTN 247

Qy 75 PDYVQSPFGRYTWERIMNFEDGAVCTVSDSSIQNC-----FIYHVKFSGLNPPNG-- 128
Db -----IVYRETSVNTTCTLHNFIFHNE-TGANENPSGVQ 288

Qy 129 --PVMQKKT-QGWEPTERLERFARDGMLIGNFMALKLEGGGHYLCEPKSTYKARKPVKMP 185
Db NIQYQTKTAQSGYNNFSLSFVYKESNFM-----YGSYHPSCNF----- 331

Qy 186 GYHYVDKLDVTNNHKDYTSVEQREISIAKPL-----VACCF 223
Db -----RLFTINGLWENSL---SVSIAYGLQGGCKQSVFSGRAATCCY 371

RESULT 14
F90891
probable enzyme [imported] - Escherichia coli (strain O157:H7, substrain R1MD 0509952)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: F90891
R;Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A;Reference number: A99629; MUID:2115231; PMID:11258796
A;Accession: F90891
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-390 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAR35525.1; PID:gl3361568; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain R1MD 0509952
C;Genetics:
A;Gene: EC82102
C;Superfamily: arylsulfatase activating enzyme atsB

Query Match 6.4%; Score 81.5; DB 2; Length 390;
Best Local Similarity 19.6%; Pred. No. 16;
Matches 55; Conservative 46; Mismatches 104; Indels 75; Gaps 12;

Qy 10 KVMMSGTVNGHYFEVEGDGKGP-----YEGQTVRLAV-TKGGPLPF 51
Db KQYIAASGNQVYFTWQ-----GEPTLAGLDFFRKVIHYQRYAGQKRIFNALQTNGLINN 105

Qy 52 AWDILSPQCQVGSIPFTKYPDIDYVQSPFGRYTW-----ERIMNFE-DGAVCTVS 103
Db EWCAFLKEHEFLVGISIDGPOELHRYRNSNGTFAKVIAAERLKSQYQVEFTLTVI 165

Qy 104 NDSSIQGNCFIYH-----VKFSG-----LNFFPN--GP 129
Db NVNVNHYPLEVYHFLKSGSKHMQFIELLEGTGNTIDFSGHSENTFRIDFSVPPTAYGK 225

Qy 130 VMQKKTQGWEPN-TERLPARDGMLIGNFMALKLEGGGHYLCEPKSTYKARKPVKMPG-- 186
Db FMSTIFMQWVKNDVGEIFIRQFESFVRFL-----GNHTSCIFQESCKDNLVVSNGDI 280

Qy 187 ---YHYVDKLDVTNNHK-DYTSVEQREISIAKPLVACC 222
Db YECDFHVPQYKLGINKSELKTNVSQLTAQKKRISAKC 320

RESULT 15

C85726
probable enzyme Z2211 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: C85726
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: C85726
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-390 <STO>
A;Cross-references: GB:AE005174; NID:gl2515177; PIDN:AAG56271.1; GSPDB:GN00145; UWGP:Z2211
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z2211
C;Superfamily: arylsulfatase activating enzyme atsB

Query Match 6.4%; Score 81.5; DB 2; Length 390;
Best Local Similarity 19.6%; Pred. No. 16;
Matches 55; Conservative 46; Mismatches 104; Indels 75; Gaps 12;

Qy 10 KVMMSGTVNGHYFEVEGDGKGP-----YEGQTVRLAV-TKGGPLPF 51
Db KQYIAASGNQVYFTWQ-----GEPTLAGLDFFRKVIHYQRYAGQKRIFNALQTNGLINN 105

Qy 52 AWDILSPQCQVGSIPFTKYPDIDYVQSPFGRYTW-----ERIMNFE-DGAVCTVS 103
Db EWCAFLKEHEFLVGISIDGPOELHRYRNSNGTFAKVIAAERLKSQYQVEFTLTVI 165

Qy 104 NDSSIQGNCFIYH-----VKFSG-----LNFFPN--GP 129
Db NVNVNHYPLEVYHFLKSGSKHMQFIELLEGTGNTIDFSGHSENTFRIDFSVPPTAYGK 225

Qy 130 VMQKKTQGWEPN-TERLPARDGMLIGNFMALKLEGGGHYLCEPKSTYKARKPVKMPG-- 186
Db FMSTIFMQWVKNDVGEIFIRQFESFVRFL-----GNHTSCIFQESCKDNLVVSNGDI 280

Qy 187 ---YHYVDKLDVTNNHK-DYTSVEQREISIAKPLVACC 222
Db YECDFHVPQYKLGINKSELKTNVSQLTAQKKRISAKC 320

Search completed: August 12, 2004, 06:13:59
Job time : 30.3012 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 06:12:47 ; Search time 14.2008 Seconds
(without alignments)
847.008 Million cell updates/sec

Title: US-09-890-463-3

Perfect score: 1268

Sequence: 1 SVIAKQMTYKYVMSTGVNH.....SIARKPIVACCFFRVKSRHK 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	187.5	14.8	238	1	GFP_AEQVI
2	87	6.9	787	1	OSTA_VIBCH
3	86.5	6.8	725	1	NCA2_MOUSE
4	86	6.8	770	1	TRPG_ASPNG
5	85	6.7	260	1	YD69_AQUAE
6	84	6.6	1162	1	VGL2_TBVB
7	83.5	6.6	385	1	YDEM_ECOLI
8	83.5	6.6	3461	1	RELN_MOUSE
9	82.5	6.5	876	1	SYA_STAAM
10	82.5	6.5	876	1	SYA_STAAM
11	82	6.5	248	1	YAL4_SCHPO
12	81.5	6.4	3462	1	RELN_RAT
13	81	6.4	1142	1	ENAM_HUMAN
14	79.5	6.3	994	1	ATA1_CHICK
15	79	6.2	403	1	H136_ARATH
16	79	6.2	687	1	CYBC_BRAVA
17	78.5	6.2	574	1	IRL2_MOUSE
18	78.5	6.2	1254	1	UBPC_YEAST
19	77.5	6.1	266	1	CB21_PETSP
20	77.5	6.1	269	1	AGAR_ECOLI
21	77.5	6.1	333	1	Y642_METUA
22	77.5	6.1	366	1	MLTA_BUCAP
23	77.5	6.1	513	1	YHCX_BACSU
24	77.5	6.1	896	1	LOXC_ARATH
25	77	6.1	726	1	RRP2_INBP9
26	76.5	6.0	1526	1	YY46_ANASP
27	76	6.0	266	1	CB24_PETSP
28	76	6.0	444	1	YGAF_ECOLI
29	76	6.0	725	1	RRP2_INBSI
30	75.5	6.0	1960	1	TF20_HUMAN
31	75	5.9	267	1	CB22_PETSP
32	75	5.9	277	1	YF99_METUA
33	75	5.9	342	1	XYNA_CALSA

ALIGNMENTS

RESULT 1
GFP_AEQVI
ID GFP_AEQVI STANDARD; PRT; 238 AA.
AC P42212; Q17104; Q27903;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Green fluorescent protein.
GN GFP.
OS Aequorea victoria (Jellyfish).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydroida; Leptomedusae;
OC Aequoreidae; Aequorea.
OX NCBI_TaxID=6100;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=92175527; PubMed=1347277;
RA Prasher D.C., Eckenrode V.K., Ward W.W., Prendergast F.G.,
RA Cormier M.J.;
RT "Primary structure of the Aequorea victoria green-fluorescent
RT protein.";
RL Gene 111:229-233(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94185810; PubMed=8137953;
RA Inouye S., Tsuji F.I.;
RT "Aequorea green fluorescent protein. Expression of the gene and
RT fluorescence characteristics of the recombinant protein.";
RL FEBS Lett. 341:277-280(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97299832; PubMed=9154981;
RA Rouwendal G.J.A., Mendes O., Wolbert E.J.H., de Boer A.D.;
RT "Enhanced expression in tobacco of the gene encoding green fluorescent
RT protein by modification of its codon usage.";
RL Plant Mol. Biol. 33:989-999(1997).
RN [4]
RP CHROMOPHORE.
RX MEDLINE=93192221; PubMed=8448132;
RA Cody C.W., Prasher D.C., Westler W.M., Prendergast F.G., Ward W.W.;
RT "Chemical structure of the hexapeptide chromophore of the Aequorea
RT green-fluorescent protein.";
RL Biochemistry 32:1212-1218(1993).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=96355665; PubMed=8703075;
RA Ormoe M., Cubitt A.B., Kallio K., Gross L.A., Tsien R.Y.,
RA Remington S.J.;
RT "Crystal structure of the Aequorea victoria green fluorescent
RT protein.";
RL Science 273:1392-1395(1996).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=98294543; PubMed=9631087;
RA Yang F., Moss L.G., Phillips G.N. Jr.;
RT "The molecular structure of green fluorescent protein.";
RL Nat. Biotechnol. 14:1246-1251(1996).

34 75 5.9 392 1 ANIA_NEIGO Q02219 neisseria g
35 75 5.9 809 1 OSTA_XANCP Q8pce0 xanthomonas
36 75 5.9 1983 1 TF20_MOUSE Q9epq8 mus musculus
37 74.5 5.9 587 1 COAT_PAVL3 P36310 parvovirus
38 74.5 5.9 2390 1 SPCP_HUMAN O15020 homo sapien
39 74 5.8 265 1 CB2B_LYCEN P07370 lycopersico
40 74 5.8 267 1 CH25_PETSP P04783 petunia sp.
41 74 5.8 398 1 Y4QJ_RHISN P55631 rhizobium s
42 74 5.8 443 1 GAT3_MOUSE P23772 mus musculus
43 74 5.8 813 1 OSTA_XANAC Q8p222 xanthomonas
44 74 5.8 876 1 SYA_STAEP Q8csa7 staphylococ
45 74 5.8 1004 1 MV10_MOUSE P23249 mus musculus


```
Db 192 PVLLPDNHYLSLSTQSALSKDPNEKRDH 217
RESULT 2
OSTA_VIBCH STANDARD; PRT; 787 AA.
AC Q9KUR9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Organic solvent tolerance protein precursor.
GN IMP OR OSTA OR VC0446.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.H., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
CC -!- FUNCTION: Determines N-hexane tolerance. Involved in outer
CC membrane permeability. Essential for envelope biogenesis. Could be
CC part of a targeting/usher system for outer membrane components (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Outer membrane (By similarity).
CC -!- SIMILARITY: Belongs to the imp/osta family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE004131; AAF93619.1; -.
DR PIR; E82323; E82323.
DR TIGR; VC0446; -.
DR HAMAP; MF_01411; -.
DR InterPro; IPR005653; Osta_C.
DR InterPro; IPR007543; Osta_C.
DR Pfam; PF03968; Osta_1.
DR Pfam; PF04453; Osta_C; 1.
KW Outer membrane; Signal; Complete proteome.
FT SIGNAL 1 28
FT CHAIN 29 787 ORGANIC SOLVENT TOLERANCE PROTEIN.
SQ SEQUENCE 787 AA; 89017 MW; 036718F1896E0F7D CRC64;
Query Match 6.9%; Score 87; DB 1; Length 787;
Best Local Similarity 19.2%; Pred. NO. 4.4;
Matches 45; Conservative 33; Mismatches 92; Indels 64; Gaps 9;
QY 6 QMTYKVMVSGTVN-----GHYFEVGGKGVGEQTVRLAVTKGGLPFA--WDIL 56
Db 398 QLSVNYAPETWKYLDLVLVSHVSFFTDARKP-----SATRVHLEPGGLKIPFSNTGNW 443
QY 57 SPQCYGSGIPFTKYPIDPVYKQSPGPRGYRIMNPFEDGAVCTVSDNSISQGNCFIYH 116
Db 444 TTEAR---VLGYTQQDLKTTD-----AKLEESVTRVPIRSV-----480
QY 117 VKFSLGNFPNGPNVQKTKQGWPNTERLFA-----RDGMLIGNFMAL-----160
Db 481 ---AGIVLERDVLDDVTQTLEPKIQYLVPEKYQDNIGLYDSTLLQTDYYGLFRSKY 537
```

```
QY 161 ----KLEGG-----CHYLCEFKSTYKAKPKVMPGHHYVDRKLDVTHNKDYTS 205
Db 538 SGVDRIESANQVSYGASTRFFDSNYKERLNIAGQIFLYDSKLINFSNKNPPOSTS 591
RESULT 3
NCA2_MOUSE
ID NCA2_MOUSE STANDARD; PRT; 725 AA.
AC P13594; Q61950;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neural cell adhesion molecule 1, 120 kDa isoform precursor (N-CAM 120)
DE (NCAM-120).
GN NCAM1 OR NCAM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6;
RX MEDLINE=87246524; PubMed=3595563;
RA Barthels D., Santoni M.-J., Wille W., Ruppert C., Caix J.-C.,
RA Hirsch M.-R., Fontecilla-Camps J.-C., Goridis C.;
RT "Isolation and nucleotide sequence of mouse NCAM cDNA that codes for
RT a Mr 79,000 polypeptide without a membrane-spanning region.";
RL EMBO J. 6:907-914(1987).
RN [2]
SEQUENCE OF 20-700 FROM N.A.
RP STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=89251563; PubMed=2721486;
RA Santoni M.-J., Barthels D., Vopper G., Boned A., Goridis C., Wille M.;
RT "Differential exon usage involving an unusual splicing mechanism
RT generates at least eight types of NCAM cDNA in mouse brain.";
RL EMBO J. 8:385-392(1989).
RN [3]
SEQUENCE OF 642-725 FROM N.A.
RX MEDLINE=88283628; PubMed=3396534;
RA Barbas J.A., Chaix J.C., Steinmetz M., Goridis C.;
RT "Differential splicing and alternative polyadenylation generates
RT distinct NCAM transcripts and proteins in the mouse.";
RL EMBO J. 7:625-632(1988).
RN [4]
SEQUENCE OF 20-36.
RP MEDLINE=86140120; PubMed=3512556;
RA Rougon G., Marshak D.R.;
RT "Structural and immunological characterization of the amino-terminal
RT domain of mammalian neural cell adhesion molecules.";
RL J. Biol. Chem. 261:3396-3401(1986).
CC -!- FUNCTION: This protein is a cell adhesion molecule involved in
CC neuron-neuron adhesion, neurite fasciculation, outgrowth of
CC neurites, etc.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=N-CAM 120;
CC IsoId=PI3594-1; Sequence=Displayed;
CC Name=N-CAM 180;
CC IsoId=PI3595-1; Sequence=External;
CC Name=N-CAM 140;
CC IsoId=PI3595-2; Sequence=External;
CC -!- SIMILARITY: Contains 5 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 2 fibronectin type III domains.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----Y00051; CAA68263.1; -
DR EMBL; X15049; CAA33148.1; ALT_SEQ.
DR EMBL; X07195; CAA30173.1; -.
DR PIR; A29673; IJWSNG.
DR PDB; 2NCM; 12-MAR-97.
DR PDB; 3NCM; 23-JUL-99.
DR MGD; MGI:97281; Ncam1.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR007110; FN_III.
DR InterPro; IPR003598; IG_c2.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00047; ig; 5.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00408; IGC2; 5.
DR PROSITE; PS00835; IG-LIKE; 5.
DR Cell adhesion; Glycoprotein; Repeat; Alternative splicing;
KW Immunoglobulin domain; Signal; Heparin-binding; GPI-anchor;
KW 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 725
FT -----NEURAL CELL ADHESION MOLECULE 1, 120 kDa
FT ISOFORM.
FT DOMAIN 20 111
FT DOMAIN 116 205
FT DOMAIN 212 302
FT DOMAIN 309 402
FT DOMAIN 407 492
FT DOMAIN 519 596
FT DOMAIN 625 692
FT DOMAIN 152 156
FT DOMAIN 161 165
FT DISULFID 41 96
FT DISULFID 139 189
FT DISULFID 235 288
FT DISULFID 330 386
FT DISULFID 427 480
FT CARBOHYD 222 222
FT CARBOHYD 316 316
FT CARBOHYD 348 348
FT CARBOHYD 424 424
FT CARBOHYD 450 450
FT CARBOHYD 479 479
FT CONFLICT 261 268
FT CONFLICT 273 273
FT CONFLICT 354 355
FT CONFLICT 549 549
FT CONFLICT 572 572
FT CONFLICT 575 575
FT CONFLICT 589 594
FT CONFLICT 600 602
FT CONFLICT 657 657
FT SEQUENCE 725 AA; 80296 MW; C2AEBB8B4461C6B2F CRC64;
Query Match 6.8%; Score 86.5; DB 1; Length 725;
Best Local Similarity 21.7%; Pred. No. 4.4;
Matches 53; Conservative 41; Mismatches 89; Indels 61; Gaps 15;
Qy 10 KVMKSGTVNGHYFVEGDKGK-PYE-----GEQTVR-LAVTKGGLPFANDILSP- 58
Db 455 KII--NTFSASYLVTPDSENDGNYNTAVNRIGQESLEFILVQADTPSPSIDRVEPY 512
Qy 59 -----QCOY-----GSIPFTKYPIDPYVKQSPFGRYTWIRIMFEDGACVTVSNDSSI 108
Db 513 SSTAQVQDEPEATGVPILKYAKBWKSLGESW-HFTWYD-----AKBANM 558
Qy 109 QNCHEI-----YHVKESGLNFPDGVNQ---KKTQG-----WEPNTERLFPADGMLI 153
Db 559 EGIVTIMGKLPETYSRLAALNGKLGELINQPSSEKTPQVPFELSAPKLEGMQMGDGNIS 618
Qy 154 GNMFMALKLEGGG---HVLCEFKSTYKARX-VKMP---GYHYVDKLDLVNHNKDYTSV 206
Db 619 KYNL--IKQDGGSPIRHYLVKYRALASEWKEPKRLPSGSHVMLKSLDWANEYEVVVA 676
```

```
Qy 207 EQRE 210
Db 677 ENQQ 680
RESULT 4
TRPG ASPNG STANDARD; PRT; 770 AA.
AC P05328;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Anthranilate synthase component II (EC 4.1.3.27) [Includes: Glutamine
DE amidotransferase; Indole-3-glycerol phosphate synthase (EC 4.1.1.48)
DE (IGPS); N-(5'-phosphoribosyl)anthranilate isomerase (EC 5.3.1.24)
DE (PRAI)].
DE TRPC.
GN TrpC.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=401;
RC MEDLINE=88223483; PubMed=2836085;
RA Kos T., Kuijvenhoven A., Hensing H.G.M., Pouwels P.H.,
RA van den Hondel C.A.M.J.J.;
RT "Nucleotide sequence of the Aspergillus niger trpC gene: structural
RT relationship with analogous genes of other organisms.";
RL Curr. Genet. 13:137-144(1988).
[2]
RN SEQUENCE OF 1-69 AND 392-433 FROM N.A.
RP MEDLINE=8613791; PubMed=2936650;
RA Kos A., Kuijvenhoven J., Wernars K., Bos C.J., van den Broek H.W.J.,
RA Pouwels P.H., van den Hondel C.A.M.J.J.;
RT "Isolation and characterization of the Aspergillus niger trpC gene.";
RL Gene 39:231-238(1985).
[1- FUNCTION: TRIFUNCTIONAL ENZYME BEARING THE GLN AMIDOTRANSFERASE
(GATASE) DOMAIN OF ANTHRANILATE SYNTHASE, INDOLE-GLYCEROLPHOSPHATE
SYNTHASE, AND PHOSPHORIBOSYLANTHRANILATE ISOMERASE ACTIVITIES.
[2- CATALYTIC ACTIVITY: N-(5-phospho-beta-D-riboseyl)-anthranilate = 1-
(2-carboxyphenylamino)-1'-deoxy-D-ribose 5-phosphate.
[3- CATALYTIC ACTIVITY: 1-(2-carboxyphenylamino)-1'-deoxy-D-ribose 5-
phosphate = 1-(indol-3-yl)glycerol 3-phosphate + CO(2) + H(2)O.
[4- CATALYTIC ACTIVITY: Chorismate + L-glutamine = anthranilate +
pyruvate + L-glutamate.
[5- PATHWAY: Tryptophan biosynthesis; first step.
[6- PATHWAY: Tryptophan biosynthesis; third step.
[7- PATHWAY: Tryptophan biosynthesis; fourth step.
[8- SIMILARITY: Contains 1 type-1 glutamine amidotransferase domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; X07071; CAA30107.1; -.
EMBL; M14403; AAA32710.1; -.
EMBL; M14404; AAA32709.1; -.
PIR; S00643; S00643.
HSP; 006129; 1QDL.
InterPro; IPR006220; Anth_synthIII.
InterPro; IPR003317; CP_synthGatase.
InterPro; IPR003009; FMN_enzyme.
InterPro; IPR000991; Gatase_1.
InterPro; IPR001468; IGPS.
InterPro; IPR001240; PRAI.
InterPro; IPR006221; TrpG_papa.
Pfam; PF00117; Gatase; 1.
```

DR Pfam; PF00218; IGPS; 1.
DR Pfam; PF00697; PRAI; 1.
DR PRINTS; PR00097; ANTSNTHASEII.
DR PRINTS; PR00099; CPSEATASE.
DR PRINTS; PR00096; GATASE.
DR PRODOM; PD001511; IGPS; 1.
DR TIGRFAMS; TIGR00566; tpgG.papA; 1.
DR PROSITE; PS00442; GATASE TYPE_I; 1.
DR PROSITE; PS00614; IGPS; 1.
KW Tryptophan biosynthesis; Isomerase; Lyase; Multifunctional enzyme;
KW Decarboxylase; Transferase; Glutamine amidotransferase.
FT DOMAIN 25 219
FT DOMAIN 255 519
FT DOMAIN 535 770
FT ACT_SITE 104 104
FT ACT_SITE 199 199
FT ACT_SITE 201 201
SQ SEQUENCE 770 AA; 82909 MW; 36D8DE5B23097012 CRC64;

Query Match 6.8%; Score 86; DB 1; Length 770;
Best Local Similarity 24.3%; Pred.No. 5.3;
Matches 50; Conservative 23; Mismatches 61; Indels 72; Gaps 12;

QY 10 KYVMGTV-NGHYFEVGDGKPYEGQTAVRLAVTKGGPLPFAWDILSPQCQVGSIPFT 68
DB 117 KVDVTGEILHGKTPKHDGK-AYEG-----LP-----GSLAVT 150
QY 69 KYPE-----DIPYVKQFFGRYTWIRIMFEDGAVCTVNSDSIQNCFTYHVKFS-- 120
DB 151 RYHSLAGTHATIPDCLEVS-----SSVQLADD-----SNKQVIMG---VRHKKLAVE 194
QY 121 GLNPENPGVQVQKTKQGWEPNTERLPFARDGMLGNFMALKI---EGGHHYLCEPKSYTK 177
DB 195 GVQFHP-----ESLTGYGRIMFRNFKLTAGTWEKGXKHFQGSSTTK 238

QY 178 ARKPVKPGHYVDRKLDVTNNHKDY 203
DB 239 ATVPSNPP--PKTDKLSILERYD 262

RESULT 5
YD69 AQUAE STANDARD; PRT; 260 AA.
AC 067381;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Hypothetical protein AQ_1369.
GN AQ_1369.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Snead M.A., Keller M., Aujaay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus.";
RL Nature 392:353-358 (1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000737; AAC07356.1; -.

DR PIR; B70419; B70419.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 260 AA; 30206 MW; B5D3FB6F37C99BB3 CRC64;

Query Match 6.7%; Score 85; DB 1; Length 260;
Best Local Similarity 22.3%; Pred.No. 1.8;
Matches 37; Conservative 28; Mismatches 71; Indels 30; Gaps 7;

QY 71 PEDIPDYVKQFFGRY-----TWIRIMFEDGAVCTVNSDSIQNCFTYHVKFSGLNPP 126
DB 46 PENVREFLKNPEKYLNIENWELQGEFD-----VQLGGNLYIVIRIPEKEPEK 97
QY 127 NGPVVQKTKQ-----GWE--PNTERLPARDGMLGNFM-ALKLEGGGHYLCE 171
DB 98 ELGIFQSVVEEAMGAFSLTALEHGWEVPKNVVIHADFVEGKNLIAAIKTEEGISTYDQ 157
QY 172 FKSTYKARKPVKPGHYVDRKLDVTNNHKD-YTSVEQREISARK 216
DB 158 LKLEEMMKVRYR--RVVYSSDVLTKYDIYDPVQSKAVIARE 201

RESULT 6
VGL2 IBVB STANDARD; PRT; 1162 AA.
ID VGL2 IBVB
AC P11223; P05134;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE E2 glycoprotein precursor (Spike glycoprotein) (Peglomer protein)
DE [Contains: Spike protein S1; Spike protein S2].
GN S.
OS Avian infectious bronchitis virus (strain Beaudette) (IBV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=111122;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85159540; PubMed=2984314;
RA Binns M.M., Boursnell M.E.G., Tomley F.M., Brown T.D.K.;
RA Brown T.D.K.;
RT "Cloning and sequencing of the gene encoding the spike protein of the
RT coronavirus IBV.";
RL J. Gen. Virol. 66:719-726 (1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87085499; PubMed=3025348;
RA Binns M.M., Boursnell M.E.G., Tomley F.M., Brown T.D.K.;
RT "Comparison of the spike precursor sequences of coronavirus IBV
RT strains M41 and 6/82 with that of IBV Beaudette.";
RL J. Gen. Virol. 67:2825-2831 (1986).
CC -!- FUNCTION: THE PEGLOMER PROTEIN MEDIATES THE BINDING OF VIRIONS
CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M95169; AAA70235.1; -.
DR EMBL; X02342; CAA26201.1; -.
DR PIR; S14939; S14939.
DR InterPro; IPR002551; Corona_S1.
DR InterPro; IPR002552; Corona_S2.
DR Pfam; PF01600; Corona_S1; 1.
DR Pfam; PF01601; Corona_S2; 1.
KW Glycoprotein; Envelope protein; Transmembrane; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1162 E2 GLYCOPROTEIN.
FT CHAIN 19 537 SPIKE PROTEIN S1.
FT CHAIN 538 1162 SPIKE PROTEIN S2.

```

FT DOMAIN 1120 1137 CYS-RICH.
FT CARBOHYD 51 51 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 103 103 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 178 178 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 247 247 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 264 264 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 425 425 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 447 447 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 513 513 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 530 530 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 579 579 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 591 591 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 676 676 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 714 714 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 947 947 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 960 960 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 979 979 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1014 1014 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1038 1038 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1051 1051 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1162 AA; 128046 MW; 0BAAD58113C8BBD5 CRC64;

Query Match 6.6%; Score 84; DB 1; Length 1162;
Best Local Similarity 21.1%; Pred. No. 13;
Matches 61; Conservative 32; Mismatches 62; Indels 134; Gaps 17;

Qy 3 IAKQMTYK-----VYMSG-----TVNGHYFVEGDGKGPYEGEQTVR 40
Db 149 VAKYPTFRSPCVNNLTSLVINGDLVLTSTNETIDVTSGAVYFKA-----192
Qy 41 LAVTKGGPLPP-----AWDL-----SPQ-----COYGSIPPTK--YPEDI 74
Db 193 -----GPIITYKVMREKALAYFNGTAQDVLCDGSPRGLACQYNTGNFSDGFYPTN 247
Qy 75 PDYVQSPFGPYTWERIMNFDGAV---CTVSDSSIQNGCFIYHVKFSGLNFPNG---128
Db 248 SSLVKQKF-----IVYRENSVNTTCTLHN-----FIFHNE--TGANPNPSPGVON 289
Qy 129 -PVMOKKT-QGWENRTERLFARDGMLIGNFMALKLEGGHYLCEFKSTYKARKEVXMPG 186
Db 290 IQTYQTKTQAGSYYNFNFSFLSSFYVYKESNFM-----YGSVHPSCKF-----331
Qy 187 YHYVDRKLDVTHNKNYTSVQRREISIAKPL-----VACCF 223
Db 332 -----RLFTINGLWNSL-----SVSIAYGFLQGGCKSVFKGRATCCY 371

```

RESULT 7

```

YDEM ECOLI
ID YDEM ECOLI STANDARD; PRT; 385 AA.
AC P76134; P77755;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ydem.
GN YDEM OR B1497.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;

```

```

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474 (1997).
RP [2]
RP SEQUENCE FROM N.A.
RX STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA Motomura K., Nakade S., Nakamura Y., Shimoto H., Nishio Y.,
RA Oshima T., Saito N., Sampei G., Seki Y., Sivasubram S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horiuchi T.;
RT "A 570-Kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377 (1996).
CC -1- SIMILARITY: BELONGS TO THE ASLB/ATSB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000247; AAC74570.1; ALT_INIT.
DR EMBL; D90791; BAA15168.1; -.
DR EMBL; D90792; BAA15171.1; -.
DR EcoGene; EG13795; ydem.
DR InterPro; IPE007197; Radical SAM.
DR Pfam; PF04055; Radical SAM; 1.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 385 AA; 44518 MW; 964E34F73E680329 CRC64;

Query Match 6.6%; Score 83.5; DB 1; Length 385;
Best Local Similarity 19.6%; Pred. No. 3.9;
Matches 55; Conservative 46; Mismatches 104; Indels 75; Gaps 12;

Qy 10 KVMYSGTVNGHYFVEGDGKGP-----YEGQTVRLAV-TKGGPLPF 51
Db 44 KQYIAASGNQVYFTWQ-----GEPTLAGLDFRKHVIHQYQYAGQKRFNALQTNGLINN 100
Qy 52 AWDILSPQCQYGSIPFTKYPEDIDYVKQSPGRTW-----ERIMNFE-DGAVCTVS 103
Db 101 EWCAFLKEHEFLVGISIDGQELHRYRNSNGTFAKVIATIERLUKSYQVFNLTVI 160
Qy 104 NDSISQGNCFYH-----VKFSG-----LNFPN--GP 129
Db 161 NVNVNHYPLEVYHFKLSIGSKHMQFIELLEGTNIDFSGHSENTFRIDFSVPTAGK 220
Qy 130 VMQKKTQGWEPN-TERLFARDGMLIGNFMALKLEGGHYLCEFKSTYKARKEVXMPG-- 186
Db 221 FMSTIFMQWKNVDVGEIFIQFESFVSRL-----GNGHTSCIFQESCKNLVVESNGDI 275
Qy 187 ---YHYVDRKLDVTHNKN-DYTSVQREISIAKPLVACC 222
Db 276 YECDFHYVPQYKIGNINKSELKTNVSVQLTAQKRIIPAKC 315

RESULT 8
RELN MOUSE
ID RELN MOUSE STANDARD; PRT; 3461 AA.
AC Q60841; Q9CUA6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Reelin precursor (EC 3.4.21.-) (Reeler protein).
OS RELN OR RL.

```

OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Cerebellum;
RX MEDLINE=95231649; PubMed=7715726;
RA D'Arcangelo G., Miao G.G., Chen S.-C., Soares H.D., Morgan J.I.,
RA Curran T.;
RT "A protein related to extracellular matrix proteins deleted in the
RT mouse mutant reeler.";
RL Nature 374:719-723(1995).
[2]
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RX MEDLINE=98086481; PubMed=9417911;
RA Royaux I., Lambert de Rouvroit C., D'Arcangelo G., Demirov D.,
RA Goffinet A.M.;
RT "Genomic organization of the mouse reelin gene.";
RL Genomics 46:240-250(1997).
[3]
RP SEQUENCE OF 2152-3461 FROM N.A. (ISOFORM 1).
RC STRAIN=BAUB/c; TISSUE=Brain;
RX MEDLINE=95375789; PubMed=7647795;
RA Hirotsune S., Takahara T., Sasaki N., Hirose K., Yoshiki A.,
RA Ohashi T., Kusakabe M., Murakami Y., Muramatsu M., Watanabe S.,
RA Nakao K., Katsuki M., Hayaishiraki Y.;
RT "The reeler gene encodes a protein with an EGF-like motif expressed by
RT pioneer neurons.";
RL Nat. Genet. 10:77-83(1995).
[4]
RP SEQUENCE OF 3044-3461 FROM N.A. (ISOFORM 2).
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glessi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staibii F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Holtmann M., Hume D.A., Kamlya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[5]
RP CHARACTERIZATION.
RX MEDLINE=97141547; PubMed=8987733;
RA D'Arcangelo G., Nakajima K., Miyata T., Ogawa M., Mikoshiba K.,
RA Curran T.;
RT "Reelin is a secreted glycoprotein recognized by the CR-50 monoclonal
RT antibody.";
RL J. Neurosci. 17:23-31(1997).
[6]
RP CHARACTERIZATION.
RX MEDLINE=21634904; PubMed=11689558;
RA Quattrocchi C.C., Wannen F., Persico A.M., Ciafre S.A.,
RA D'Arcangelo G., Farace M.G., Keller F.;
RT "Reelin is a serine protease of the extracellular matrix.";
RL J. Biol. Chem. 277:303-309(2002).
[7]
RP TISSUE SPECIFICITY.
RX MEDLINE=97325946; PubMed=9182958;
RA Schiffmann S.N., Bernier B., Goffinet A.M.;
RT "Reelin mRNA expression during mouse brain development.";
RL Eur. J. Neurosci. 9:1055-1071(1997).
[8]
RP ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RX MEDLINE=99263436; PubMed=10328932;
RA Lambert de Rouvroit C., Bernier B., Royaux I., de Bergueyck V.,
RA Goffinet A.M.;
RT "Evolutionarily conserved, alternative splicing of reelin during brain
RT development.";
RL Exp. Neurol. 156:229-238(1999).
[9]
RP BINDING TO VLDLR AND ApoER2.
RX MEDLINE=20036019; PubMed=10571241;
RA Hiesberger T., Trommsdorff M., Howell B.W., Goffinet A.M., Mumby M.C.,
RA Cooper J.A., Herz J.;
RT "Direct binding of Reelin to VLDL receptor and ApoE receptor 2 induces
RT tyrosine phosphorylation of disabled-1 and modulates tau
RT phosphorylation.";
RL Neuron 24:481-489(1999).
[10]
RP FUNCTION.
RX MEDLINE=20359755; PubMed=10880573;
RA Yip J.W., Yip Y.P.L., Nakajima K., Capriotti C.;
RT "Reelin controls position of autonomic neurons in the spinal cord.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:8612-8616(2000).
CC -!- FUNCTION: Extracellular matrix serine protease that plays a role
CC in layering of neurons in the cerebral cortex and cerebellum.
CC Regulates microtubule function in neurons and neuronal migration.
CC Affects migration of sympathetic preganglionic neurons in the
CC spinal cord, where it seems to act as a barrier to neuronal
CC migration. Enzymatic activity is important for the modulation of
CC cell adhesion. Binding to the extracellular domains of lipoprotein
CC receptors VLDLR and ApoER2 induces tyrosine phosphorylation of
CC Dab1 and modulation of tau phosphorylation.
CC -!- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q60841-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q60841-2; Sequence=VSP_005577;
CC Name=3;
CC IsoId=Q60841-3; Sequence=VSP_005578;
CC -!- TISSUE SPECIFICITY: The major isoform 1 is neuron-specific. It is
CC abundantly produced during brain ontogenesis by the Cajal-Retzius
CC cells and other pioneer neurons located in the telencephalic
CC marginal zone and by granule cells of the external granular layer
CC of the cerebellum. Expression is located in deeper layers in the
CC developing hippocampus and olfactory bulb, low levels of
CC expression are also detected in the immature striatum. At early
CC developmental stages, expressed also in hypothalamic
CC differentiation fields, tectum and spinal cord. A moderate to low
CC level of expression occurs in the septal area, striatal fields,
CC habenular nuclei, some thalamic nuclei, particularly the lateral
CC geniculate, the retina and some nuclei of the reticular formation
CC in the central field of the medulla. Very low levels found in
CC liver and kidney. No expression in radial glial cells, cortical
CC plate, Purkinje cells and inferior olivary neurons. The minor
CC isoform 2 is only expressed in non neuronal cells. The minor
CC isoform 3 is found in the same cells as isoform 1, but is almost
CC undetectable in retina and brain stem.
CC -!- DEVELOPMENTAL STAGE: First detected at embryonic day 11.5.
CC Expression increases up to birth and remains high from post-natal
CC day 2 to 11 in both cerebellum and fore/midbrain. Expression
CC declines thereafter and is largely brain specific in the adult.
CC -!- DOMAIN: The basic C-terminal region is essential for secretion.
CC -!- PTM: N-glycosylated and to a lesser extent also O-glycosylated.
CC -!- DISEASE: Defects in reelin are the cause of the autosomal recessive
CC reeler (rl) phenotype which is characterized by impaired motor
CC coordination, tremors and ataxia. Neurons in affected mice fail to
CC reach their correct locations in the developing brain, disrupting
CC the organization of the cerebellar and cerebral cortices and other

```
CC CC laminated regions.
CC CC -!- SIMILARITY: Belongs to the reelin family.
CC CC -!- SIMILARITY: Contains 8 EGF-like domains.
CC CC -!- SIMILARITY: Contains 15 BNR repeats.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; U24703; AAB91599.1; -.
CC CC EMBL; D63520; BAA09788.1; ALT INIT.
CC CC EMBL; AK017094; BAB30592.1; -.
CC CC MGD; MGI:103022; ReIn.
CC CC GO; GO:0005635; C:extracellular space; IDA.
CC CC GO; GO:0007420; P:brain development; IMP.
CC CC GO; GO:0016477; P:cell migration; IMP.
CC CC InterPro; IPR006209; EGF like.
CC CC InterPro; IPR002860; GH BNR.
CC CC InterPro; IPR006210; IEGF.
CC CC InterPro; IPR002861; Reeler.
CC CC Pfam; PF02012; BNR; 15.
CC CC Pfam; PF00008; EGF; 3.
CC CC Pfam; PF02014; Reeler; 1.
CC CC SMART; SM00181; EGF; 5.
CC CC PROSITE; PS00022; EGF_1; 7.
CC CC PROSITE; PS01186; EGF_2; 6.
CC CC PROSITE; PS00026; EGF_3; 5.
CC CC Hydrolase; Serine protease; Developmental protein; Matrix protein;
CC CC Cell adhesion; EGF-like domain; Glycoprotein; Repeat; Signal;
CC CC Alternative splicing.
CC CC FT SIGNAL 1 26
CC CC FT CHAIN 27 3461
CC CC FT DOMAIN 40 172
CC CC FT DOMAIN 671 702
CC CC FT DOMAIN 1030 1061
CC CC FT DOMAIN 1409 1442
CC CC FT DOMAIN 1765 1796
CC CC FT DOMAIN 2129 2161
CC CC FT DOMAIN 2478 2509
CC CC FT DOMAIN 2853 2884
CC CC FT DOMAIN 3228 3260
CC CC FT REPEAT 593 604
CC CC FT REPEAT 799 810
CC CC -----
CC CC Query Match 6.6%; Score 83.5; DB 1; Length 3461;
CC CC Best Local Similarity 20.3%; Pred. No. 53;
CC CC Matches 57; Conservative 24; Mismatches 75; Indels 125; Gaps 12;
CC CC -----
CC CC 12 YMSGTVNGHYVEVDGK---GKPYEGEQTVRLAVTKG---GFLPFAWDI---LSQP 59
CC CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC CC 2080 YYAGTTQWRREVVFHFKLHCG-----SVRFWYQGFYPASQPTWALDNVYIGPQ 2132
CC CC Qy :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC CC 60 CQ---YG-----SIPFKYKPEDIPDYVKQSPFGRYTWERIMNF----- 94
CC CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC CC 2133 CEEMCYGHGSCINGTKICDPGSGPTCKISTKNPDLKDDFEGQLESDRFLMSGGRPS 2192
CC CC Qy 95 -----EDGAVCTVNSDSIQGNCFYHYKVF-----SGLNFPNGP 129
CC CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC CC 2193 KCGILSGNNLFFNEDGLRLMLVTRDLS-----HARFVQFFMRLCGKGVDPDRSQP 2246
CC CC Qy 130 VM-----QKKTQGWEPNTERLF--- 146
CC CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC CC 2247 VLLQYSLNGLSWSLLQBFLEFSNSNVGYIALEMLPKARSGSTELRWQFSENGHFYSP 2306
CC CC Qy 147 -ARDGMLIGNFMALKLGGGHLYCEFKSTYKARKPVQMPG 186
CC CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC CC 2307 WWIDQILGGNI-----SGNTVLEDDFSLDSRKLHLHPG 2341
```

```
RESULT 9
SYA STAA
ID SYA STAA STANDARD; PRT; 876 AA.
AC Q99TNI;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Alanyl-tRNA synthetase (EC 6.1.1.7) (Alanine--tRNA ligase) (ALARS).
GN ALAS OR SAV1618 OR SA1446.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699), and
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699, and N315;
RX MEDLINE=213111952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
RT aureus."
RL Lancet 357:1225-1240(2001).
CC -!- CATALYTIC ACTIVITY: ATP + L-alanine + tRNA(Ala) = AMP +
CC diphosphate + L-alanyl-tRNA(Ala).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; AP003362; BAB57780.1; -.
CC CC EMBL; AP003134; BAB42710.1; -.
CC CC FIR; AB9944; AB9944.
CC CC HAMAP; MF_00036; -.
CC CC InterPro; IPR003156; DHHA1.
CC CC InterPro; IPR002318; tRNA-synt 2c.
CC CC InterPro; IPR006193; tRNA_synt_Ala.
CC CC Pfam; PF02272; DHHA1; 1.
CC CC Pfam; PF01411; tRNA-synt_2c; 1.
CC CC PRINTS; PR00980; TRNASYNTHALA.
CC CC TIGRFAMs; TIGR00344; alas; 1.
CC CC PROSITE; PS0860; AA tRNA LIGASE II ALA; 1.
CC CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
CC CC Complete proteome.
CC CC SEQUENCE 876 AA; 98538 MW; 2B2BC79041AC264F CRC64;
CC -----
CC CC Query Match 6.5%; Score 82.5; DB 1; Length 876;
CC CC Best Local Similarity 21.9%; Pred. No. 13;
CC CC Matches 46; Conservative 25; Mismatches 60; Indels 79; Gaps 11;
CC CC -----
CC CC 16 TVNGH-IPVEVDGKGRPYEGEQTVRLAVTKGGLPFAWDILSPQCGSIP----FTKY 70
CC CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC CC 81 TARHTTFFMLGNFGSISGDFYKQEAIE-----FAWEFLTSDKWMGMGPDKLYVTIH 130
CC CC Qy :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC CC 71 PEDIPDYVKQSPFGRYTWERIMNFEDGAVCTVNSDSIQGNCFYHYKVFSGLNFPNGPV 130
CC CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC CC 131 PEDMEAY-----NIWKIDIGLESRII-----RIEGN-----FWDIGEGPSG-- 167
CC CC Qy :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC CC 131 MQKTTQGWEPNTERLFARDGMLIGNFMALKLGGGHLYCEFKSTYKARKPVQMPGCIHYV 190
CC CC Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
CC CC 168 -----PNTEIFYDR-GEAYGQDDPABEMYPGEN----- 195
CC CC Qy 191 DRKLDV-----TNHNKD--YTSVEQREI 211
```

[illegible]

```
Db 164 VDVEGFEGLGSTDCEBQIFKPRQALTFQGHPEFSTEVVNTMKVLRGTEVFTE-QQK 222
Qy 210 EISAR 215
| : :
Db 223 BEALKR 228

RESULT 12
RELN RAT
ID RELN RAT STANDARD; PRT; 3462 AA.
AC P58751; Q80T65;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Reelin precursor (EC 3.4.21.-).
GN RELN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN NCBI_TaxID=10116;
SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Cerebellum;
RA Kikkawa S., Terashima T.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A., AND DISEASE.
RC TISSUE=Cerebellum;
RA Yokoi N., Nanae M., Wang H.-W., Kojima K., Fuse M., Yasuda K.,
RA Serikawa T., Seino S., Komeda K.;
RT "Rat neurological disease creeping is caused by a mutation in the
RL reelin gene.";
RL Brain Res. Mol. Brain Res. 112:1-7(2003).
RN [3]
ALTERNATIVE SPLICING.
RX MEDLINE=9263436; PubMed=10328932;
RA Lambert de Rouvroit C., Bernier B., Royaux I., de Bergueyck V.,
RA Goffinet A.M.;
RT "Evolutionarily conserved, alternative splicing of reelin during brain
RT development.";
RL Exp. Neurol. 156:229-238(1999).
CC -!- FUNCTION: Extracellular matrix serine protease that plays a role
CC in layering of neurons in the cerebral cortex and cerebellum.
CC Regulates microtubule function in neurons and neuronal migration.
CC Affects migration of sympathetic preganglionic neurons in the
CC spinal cord, where it seems to act as a barrier to neuronal
CC migration. Enzymatic activity is important for the modulation of
CC cell adhesion. Binding to the extracellular domains of lipoprotein
CC receptors VLDLR and ApoER2 induces tyrosine phosphorylation of
CC Dab1 and modulation of Tau phosphorylation (By similarity).
CC -!- SUBUNIT: Binds to the ectodomains of VLDLR and ApoER2 (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Name=1; IsoId=P58751-1; Sequence=Displayed;
CC Name=2; IsoId=P58751-2; Sequence=VSP_005579;
CC Name=3; IsoId=P58751-3; Sequence=VSP_005580;
CC -!- TISSUE SPECIFICITY: Abundantly produced during brain ontogenesis
CC by the Cajal-Retzius cells and other pioneer neurons located in
CC the telencephalic marginal zone and by granule cells of the
CC external granular layer of the cerebellum.
CC -!- DOMAIN: The basic C-terminal region is essential for secretion (By
CC similarity).
CC -!- DISEASE: Defects in Reelin are the cause of creeping, which is
CC characterized by tremor, gait ataxia, cerebellar hypoplasia and
CC abnormal neuronal migration (particularly in the cerebral cortex
CC and hippocampus). The mutation is due to a nucleotide insertion at
CC codon 1892 which results in a translational frameshift and
```

```
CC truncation of the protein.
CC -!- SIMILARITY: Belongs to the reelin family.
CC -!- SIMILARITY: Contains 8 EGF-like domains.
CC -!- SIMILARITY: Contains 15 BNR repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB049473; BAB78470.1; -.
DR EMBL; AB062680; BAC75467.1; -.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR002860; GH_BNR.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002861; Reeler.
DR Pfam; PF02012; BNR; 15.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF02014; Reeler; 1.
DR SMART; SM00181; EGF; 6.
DR PROSITE; PS00022; EGF_1; 7.
DR PROSITE; PS01186; EGF_2; 6.
DR PROSITE; PS50026; EGF_3; 5.
DR Hydrolase; Serine protease; Developmental protein; Matrix protein;
KW Cell adhesion; EGF-like domain; Glycoprotein; Repeat; Signal;
KW Alternative splicing.
FT SIGNAL 1 27 POTENTIAL.
FT CHAIN 28 3462 REELIN.
FT DOMAIN 41 173 REELER.
FT DOMAIN 672 703 EGF-LIKE 1.
FT DOMAIN 1031 1062 EGF-LIKE 2.
FT DOMAIN 1410 1443 EGF-LIKE 3.
FT DOMAIN 1766 1797 EGF-LIKE 4.
FT DOMAIN 2130 2162 EGF-LIKE 5.
FT DOMAIN 2479 2510 EGF-LIKE 6.
FT DOMAIN 2854 2885 EGF-LIKE 7.
FT DOMAIN 3229 3261 EGF-LIKE 8.
FT REPEAT 594 605 BNR 1.
FT REPEAT 800 811 BNR 2.
FT REPEAT 953 964 BNR 3.
FT REPEAT 1158 1169 BNR 4.
FT REPEAT 1324 1335 BNR 5.
FT REPEAT 1536 1547 BNR 6.
FT REPEAT 1687 1698 BNR 7.
FT REPEAT 1885 1896 BNR 8.
FT REPEAT 2044 2055 BNR 9.
FT REPEAT 2251 2262 BNR 10.
FT REPEAT 2400 2411 BNR 11.
FT REPEAT 2599 2610 BNR 12.
FT REPEAT 2779 2790 BNR 13.
FT REPEAT 2980 2991 BNR 14.
FT REPEAT 3364 3375 BNR 15.
FT DOMAIN 3433 3462 ARG-RICH (BASIC).
FT CARBOHYD 142 142 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 259 259 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 630 630 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1268 1268 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1448 1448 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1601 1601 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1751 1751 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1922 1922 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2146 2146 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2270 2270 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2318 2318 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2570 2570 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 2963 2963 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3017 3017 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 3074 3074 N-LINKED (GLCNAC. .) (POTENTIAL).
```

```

CC CARBOHYD 3186 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 3413 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 3440 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC VARSPLIC 3430 Missing (in isoform 2).
CC VARSPLIC 3431 /FTID=VSP 005579.
CC VARSPLIC 3432 Missing (in isoform 3).
CC VARSPLIC 3433 /FTID=VSP 005580.
CC CONFLICT 336 H -> R (IN REF. 2).
CC CONFLICT 2714 V -> L (IN REF. 2).
CC SEQUENCE 3462 AA; 387525 MW; FCCR89B0905035F6 CRC64;

Query Match 6.4%; Score 81.5; DB 1; Length 3462;
Best Local Similarity 19.9%; Pred. No. 80;
Matches 56; Conservative 25; Mismatches 75; Indels 125; Gaps 12;

QY 12 YMSGTVNGHYFEVGGDK---GKPYEGEQTVRLAVTKG---GPLPFAMDI---LSPQ 59
Dd 2081 YYAGTTQGMRRVHVHFKLHLCG-----SVRFWYQGFYAGPQVTVWADNVVIGPQ 2133
QY 60 CQ-----YGSI-----PFTKYPEDIPDYVKQSPGRTYTWERINMF----- 94
Dd 2134 CEMCCGHCSCWNGYKICIDPGYSQGTCKISKYKNDPDLKDDFEGQLESDFLMSGGKPS 2193
QY 95 -----EDGAVCTVSNDSSIQGCFIYHVKF-----SGLNFPNPGP 129
Dd 2194 RKCGILSSGNLFFNEDGLRLVTRDLDS-----HARFVQFFMRLGCGKGVDPDRSOP 2247
QY 130 VM-----OKTKQGWEPNTERLF--- 146
Dd 2248 VLLQYSLNGLSWSLLQELFFSNSSNGVRYIALEMPLKARSGSTRLRWQPSENGHFYSP 2307
QY 147 -ARDGMLIGNFMALKEGGHGLCEFKSTYKARKPVKMPG 186
Dd 2308 WVIDQLIGNI-----SGNTVLEDDFSLDRKWLHPG 2342

RESULT 13
ENAM HUMAN
ID ENAM HUMAN STANDARD; PRT; 1142 AA.
AC Q9NRML; Q9H3D1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Enamelin precursor.
CN ENAM.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Hu C.-C., Qian Q., Zhang C., Fukae M., Uchida T., Simmer J.P.;
RT "cDNA sequence of human enamel."
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1014-1142 FROM N.A.
RX MEDLINE=20489450; PubMed=11037750;
RA Dong J., Gu T.T., Simmons D., MacDougall M.;
RT "Enamelin maps to human chromosome 4q21 within the autosomal dominant
RT amelogenesis imperfecta locus."
RL Eur. J. Oral Sci. 108:353-358(2000).
CC -!- FUNCTION: Involved in the mineralization and structural
CC organization of enamel. Involved in the extension of
CC the secretory stage of dental enamel formation.
CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -!- TISSUE SPECIFICITY: Expressed in tooth particularly in
CC odontoblast, ameloblast and cementoblast.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF125373; AAG43242.1; -.
CC EMBL; AF210247; AAF73847.1; -.
CC Genew; HGNC:3344; ENAM.
CC MIM; 606585; -.
CC DR GO:0005578; C:extracellular matrix; NAS.
CC DR GO:0030345; F:structural constituent of tooth enamel; NAS.
CC DR GO:0030282; P:bone mineralization; NAS.
CC DR GO:0042476; P:odontogenesis; NAS.
CC KW Biomineralization; Extracellular matrix; Glycoprotein; Signal.
FT SIGNAL 1 39 POTENTIAL.
FT CHAIN 40 1142 ENAMELIN.
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 467 467 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 534 534 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 934 934 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1040 1040 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1142 AA; 128745 MW; 77419C4375EAD6EC CRC64;

Query Match 6.4%; Score 81; DB 1; Length 1142;
Best Local Similarity 29.5%; Pred. No. 24;
Matches 23; Conservative 12; Mismatches 35; Indels 8; Gaps 3;

QY 18 NGHYFEVGGDKPYEGEQTVRLAVTKG---GPLPFAMDI---GKPYEGEQTVRK 72
Dd 493 NSYI---PRGDSRKVPNSDQTSQNLPGKIVLGSRMPYESETNOSLKHSSYQAVIPE 550
QY 73 DIPDYVKQSPF---GRYTWE 89
Dd 551 EIPSPAKEHFFPAGRNWTD 568

RESULT 14
ATATL CHICK
ID ATATL CHICK STANDARD; PRT; 994 AA.
AC P13585;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Sarcoplasmic/endoplasmic reticulum calcium ATPase 1 (EC 3.6.3.8)
DE (Calcium pump 1) (SERCAL) (SR Ca(2+)-ATPase 1) (Calcium-transporting
DE ATPase sarcoplasmic reticulum type, fast twitch skeletal muscle
DE isoform) (Endoplasmic reticulum class 1/2 Ca(2+) ATPase).
CN ATP2A1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89313743; PubMed=2526293;
RA Karin N.J., Kaprielian Z., Fambrough D.M.;
RT "Expression of avian Ca2+-ATPase in cultured mouse myogenic cells.";
RL Mol. Cell. Biol. 9:1978-1986(1989).
RN [2]
RP REVISIONS.
RA Karin N.J.;
RL Submitted (XXX-1991) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: This magnesium-dependent enzyme catalyzes the hydrolysis
CC of ATP coupled with the translocation of calcium from the cytosol
CC to the sarcoplasmic reticulum lumen. Contributes to calcium
CC sequestration involved in muscular excitation/contraction (by
CC similarity).
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + Ca(2+) (Cis) = ADP + phosphate +
CC Ca(2+) (Trans).
```


RA Chan M.M., Tang C.C., Onodera C.S., Deng J.M., Akiyama K., Ansari Y.,
RA Atakawa T., Banh J., Banno F., Bowser L., Brooks S.Y., Carninci P.,
RA Chao Q., Choy N., Enju A., Goldsmith A.D., Gurjal M., Hansen N.F.,
RA Hayashizaki Y., Johnson-Hopson C., Hsuan V.W., Iida K., Karnes M.,
RA Khan S., Koeseema E., Ishida J., Jiang P.X., Jones T., Kawai J.,
RA Kamiya A., Meyers C., Nakajima M., Narusaka M., Seki M., Sakurai T.,
RA Satou M., Tanse R., Vaysberg M., Wallender E.K., Wong C., Yamamura Y.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,
RT "Empirical analysis of transcriptional activity in the Arabidopsis
genome.";
RL Science 302:842-846(2003).
CC -!- FUNCTION: Essential for photosystem II (PSII) biogenesis; putative
CC protein stability or assembly factor for PSII.
CC -!- SUBCELLULAR LOCATION: Chloroplast; within the thylakoid lumen but
CC attached to the membrane. Restricted to the stromal lamellae.
CC -!- TISSUE SPECIFICITY: Expression in green tissue, not roots.
CC -!- DEVELOPMENTAL STAGE: Accumulates also in dark-grown seedlings.
CC -!- SIMILARITY: Belongs to the ycf48 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y15628; CAA75723.1; -.
DR EMBL; AB006708; BAB09829.1; -.
DR EMBL; AY045691; AAK74049.1; -.
DR PIR; T51828; T51828.
DR InterPro; IPR002860; GH_BNR.
DR InterPro; IPR006311; Tat.
DR Pfam; PF02012; ENR; 4.
DR TIGRFAMs; TIGR01409; Tat signal seq; 1.
DR Chloroplast; Thylakoid; Transit peptide; Photosystem II; Membrane.
KW TRANSIT 1 53 CHLOROPLAST (POTENTIAL).
FT TRANSIT 54 78 THYLAKOID (POTENTIAL).
FT CHAIN 79 403 PHOTOSYSTEM II STABILITY/ASSEMBLY FACTOR
FT HCF136.
SQ SEQUENCE 403 AA; 44103 MW; 11079552F817FF9D CRC64;

Query Match 6.2%; Score 79; DB 1; Length 403;
Best Local Similarity 23.9%; Pred. No. 10;
Matches 49; Conservative 25; Mismatches 89; Indels 42; Gaps 11;

Qy 18 NGHVFVEGDGKGYEGEQTIVRLAVTKGGLPFAWDILSPQCQY-GSIPFTKYPED--- 73
Db 144 NYRENSISFKGKGWIIKGPAILLYTADAGE---NWDRIPLSSQLPGDMVFIKATEDKSA 200
Qy 74 --IPD---YVKQSPFGYTWERIMNFDGAVCTVSDSSIQGNCFIYHVKESGLNFPFN 127
Db 201 EMVTDGAIYVTSN--RGYNKAAIQETVSATLNTVSSGISGASY-YTGTFSAVNRSPD 257
Qy 128 GPVMQKTKQG-----WEFNT-----ERLFARDGMLIG---NNFMALKEGGHYLC--- 170
Db 258 GRVAVSSRGNGFFLTWEGQPYWQPHNRAVARIRIQNWGWRADGGLWLLVIRGGGLYLSKGT 317
Qy 171 ----EFKSTYKARKPKVMPGHYVD 191
Db 318 GITBEFEFV-----PVQSRGFGILD 337

Search completed: August 12, 2004, 06:20:08
Job time : 16.2008 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 06:12:47 ; Search time 60.1168 Seconds

(without alignments)
1212.385 Million cell updates/sec

Title: US-09-890-463-3

Perfect score: 1268

Sequence: 1 SVIAKQMTYKYVMSGTVNGH.....SIARKPLVACFFPRVKSRHK 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

1: sp archaea:*

2: sp bacteria:*

3: sp fungi:*

4: sp human:*

5: sp invertebrate:*

6: sp mammal:*

7: sp mhc:*

8: sp organelle:*

9: sp phage:*

10: sp plant:*

11: sp rodent:*

12: sp virus:*

13: sp vertebrate:*

14: sp unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archheap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1138	89.7	221	5 Q95P04	Q95P04 gonipora t
2	780	61.5	225	5 Q90Y8	Q90Y8 discosoma s
3	750	59.1	230	5 Q9GTJ7	Q9GTJ7 discosoma s
4	737.5	58.2	232	5 Q9U6Y7	Q9U6Y7 discosoma s
5	704.5	55.6	236	5 Q8T6U0	Q8T6U0 dendroneph
6	703	55.4	225	5 Q963F5	Q963F5 montastraea
7	698	55.0	225	5 Q7Z0W4	Q7Z0W4 montastraea
8	672	53.0	225	5 Q7Z0W5	Q7Z0W5 montastraea
9	670	52.8	266	5 Q9U6Y3	Q9U6Y3 clavularia
10	667.5	52.6	227	5 Q7Z0W6	Q7Z0W6 montastraea
11	665.5	52.5	225	5 Q7Z0W9	Q7Z0W9 montastraea
12	665.5	52.5	227	5 Q962P9	Q962P9 montastraea
13	665.5	52.5	227	5 Q7Z0W8	Q7Z0W8 montastraea
14	664	52.4	225	5 Q95U07	Q95U07 montastraea
15	655.5	51.7	234	5 Q8T5F2	Q8T5F2 montastraea
16	654.5	51.6	225	5 Q8T5F1	Q8T5F1 montastraea

17	654.5	51.6	234	5	Q7Z0W7	Q7Z0W7 montastraea
18	652	51.4	224	5	Q8MU48	Q8MU48 montastraea
19	651.5	51.4	227	5	Q95V10	Q95V10 montastraea
20	636.5	50.2	225	5	Q8I6J8	Q8I6J8 trachyphyl
21	623.5	49.2	234	5	Q8MU47	Q8MU47 montastraea
22	612.5	48.3	259	5	Q8MMA2	Q8MMA2 agaricia fr
23	607	47.9	231	5	Q8T5B9	Q8T5B9 ricordea fl
24	600.5	47.4	231	5	Q8T6T8	Q8T6T8 discosoma s
25	600	47.3	231	5	Q8T5B8	Q8T5B8 ricordea fl
26	586.5	46.3	231	5	Q8I5F8	Q8I5F8 parascyoni
27	584.5	46.1	239	5	Q8MMA1	Q8MMA1 agaricia ag
28	578	45.6	231	5	Q8MU46	Q8MU46 ricordea fl
29	571.5	45.1	232	5	Q9GP15	Q9GP15 anemonia su
30	566.5	44.7	227	5	Q95W86	Q95W86 condylactis
31	562.5	44.4	227	5	Q95W85	Q95W85 radianthus
32	561	44.2	228	5	Q9GP16	Q9GP16 anemonia su
33	558.5	44.0	227	5	Q95W11	Q95W11 condylactis
34	558.5	44.0	232	5	Q9GZ28	Q9GZ28 anemonia su
35	546.5	43.1	227	5	Q8MU45	Q8MU45 condylactis
36	537	42.4	228	5	Q86LV4	Q86LV4 radianthus
37	535	42.2	214	5	Q86LV7	Q86LV7 meandrina m
38	534.5	42.2	235	5	Q8T5F0	Q8T5F0 scolymia cu
39	530.5	41.8	234	5	Q8T5F3	Q8T5F3 scolymia cu
40	530	41.8	214	5	Q86LV8	Q86LV8 meandrina m
41	529.5	41.8	225	5	Q8T6T9	Q8T6T9 radianthus
42	528	41.6	229	5	Q9U6Y6	Q9U6Y6 anemonia ma
43	519.5	41.0	229	5	Q8T5E7	Q8T5E7 condylactis
44	485.5	38.3	231	5	Q9U6Y5	Q9U6Y5 zoanthus sp
45	482.5	38.1	238	5	Q9BLY9	Q9BLY9 renillia mue

ALIGNMENTS

RESULT 1

Q95P04 ID Q95P04 PRELIMINARY; PRT: 221 AA.
AC Q95P04;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE GFP-like chromoprotein.
OS Gonipora tenuidens.
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Fungi; Poritidae; Gonipora.
OX NCBI_TaxID=75301;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21538626; PubMed=11682051;
RA Gurskaya N.G., Fradkov A.F., Tersikh A., Matz M.V., Labas Y.A.,
EA Martynov V.I., Yanushevich Y.G., Lukyanov K.A., Lukyanov S.A.;
RT "GFP-like chromoproteins as a source of far-red fluorescent
proteins(1).";
RL FEBS Lett. 507:16-20(2001).
DR EMBL; AF383156; AAL27542.1; -.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR ProDom; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 221 AA; 24918 MW; 93F9F4B5C2003CB4 CRC64;

Query Match 89.7%; Score 1138; DB 5; Length 221;
Best Local Similarity 95.0%; Pred. No. 5e-100;
Matches 209; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
QY 1 SVIAKQMTYKYVMSGTVNGHYFEVGDGKGPYEQTVRLAVTKGGLPFPFADWILSPQC 60
Db 2 SVIAKQMTYKYVMSGTVNGHYFEVGDGKGPYEQTVRLAVTKGGLPFPFADWILSPQS 61
QY 61 QYGSIPFTKYPEDIPDYVKQSPFGYTWERIMNFDGAVCTVSDSSIQGNCFIYVVKFS 120
Db 62 QYGSIPFTKYPEDIPDYVKQSPFGYTWERIMNFDGAVCTVSDSSIQGNCFIYVVKFS 121

Q063F5;
AC 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescent protein.
DE Montastraea cavernosa (great star coral).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Faviina; Faviidae; Montastrea.
OX NCBI_TaxID=63558;
RN [1]
RN SEQUENCE FROM N.A.
RP Lesser M.P., Barry T.M., Mazel C., Matz M.V., Lukyanov S.A.,
RA Falkowski P., Gorbunov M., Kolber Z.;
RT "Green fluorescent proteins in Caribbean Scleractinian corals.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF384683; AAK62982.2; -;
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP-like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFPLORESCENT.
DR ProDom; PD013756; Green_fl_protein; 1.
DR SEQUENCE 225 AA; 25847 MW; 77DE7D7C616929AF CRC64;

Query Match 55.4%; Score 703; DB 5; Length 225;
Best Local Similarity 56.3%; Pred. No. 1.1e-58;
Matches 121; Conservative 46; Mismatches 48; Indels 0; Gaps 0

QY 1 SVIAKQMTYKYVMGSGTVNGHYFEVGDGKPKYEGEQTIVRLAVTKGGPLPFAWDILSPQC 60
DB 2 SVIKPIMEIKLRMQGVVNGHKFKVKGEGKPFEGTQTINLTVKEGAPLPFAWDILTSAF 61

QY 61 QYGSIPFTKYPEDIPDVKGSFPGRYTWERIMNFDGAVCTVSDSSIQGNCFTIYHVKFS 120
DB 62 QYGNRVFTKYPDDIPDYFKQTFPEGYSWERIMAYEDOSICATSDIKMEGDCFTIYEQFH 121

QY 121 GLNPPNGPVMQKKTQGWENTELFARDGLIGNFMALKEGGHYLCBFKSTYKARK 180
DB 122 GVNPPNGPVMQKKTLEWPESTEKMYVRDGVKGDVNMALLLEGGHYRCDFRSYKAKK 181

QY 181 PVKMPGHHYVDKLDVTNNHKNKYTSVEQREISAR 215
DB 182 RVQLPDYHFVHDRIEILSHDNDYNTVKLSDEAEAR 216

RESULT 7
Q7Z0W4 PRELIMINARY; PRT; 225 AA.
ID Q7Z0W4 AC Q7Z0W4;
AC Q7Z0W4; (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescent protein.
DE Montastraea cavernosa (great star coral).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Faviina; Faviidae; Montastrea.
OX NCBI_TaxID=63558;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=mc6;
RC MEDLINE=2269801; PubMed=12777529;
RX Kelmanson I.V., Matz M.V.;
RT "Molecular Basis and Evolutionary Origins of Color Diversity in Great
RT Star Coral Montastrea cavernosa (Scleractinia: Faviida).";
RL Mol. Biol. Evol. 20:1125-1133 (2003).
DR EMBL; AY181557; AAO61603.1; -;
DR SEQUENCE 225 AA; 25827 MW; A600ADD716C5921E CRC64;

Query Match 55.0%; Score 698; DB 5; Length 225;
Best Local Similarity 55.8%; Pred. No. 3.3e-58;
Matches 120; Conservative 43; Mismatches 52; Indels 0; Gaps 0

QY 1 SVIAKQMTYKYVMGSGTVNGHYFEVGDGKPKYEGEQTIVRLAVTKGGPLPFAWDILSPQC 60

```
Db 2 SVIKPDMKIKLRMEGAVNGHNFVIEGEGKGFEGTQTINLTVRKGGPLPFAYDILTAA 61
Qy 61 QYGSIPFTKYPEDIPDYVKQSPGRTYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKS 120
Db 62 QYGNRAFTKYPDIADYFKQSPGYSWERSMTYEDQGICIIKSIDRMEGDCFIYEIRYD 121
Qy 121 GLNFPNGPVQKKTQGGNEPNTERLFARDGMLIGNNFMALKLEGGHYLCFFKSTYKARK 180
Db 122 GVNFPSSGPVMQKTKLWEPSTKMYRVDGVLKGDVNMALLLEGGHYRCDFRSTYKAKK 181
Qy 181 PVKMPGYHYVDRKLDVTNNHNDYTSVEQREISIA 215
Db 182 RVQLPDYHFVDRHRIELSHDNDYNTVKLSNAEAR 216
```

```
RESULT 8
Q7ZOW5 PRELIMINARY; PRT; 225 AA.
AC Q7ZOW5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Cyan fluorescent protein.
OS Montastraea cavernosa (great star coral).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Faviina; Faviidae; Montastraea.
OX NCBI_TaxID=63558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mc5;
RX MEDLINE=22689801; PubMed=12777529;
RA Kelmanson I.V., Matz M.V.;
RT "Molecular Basis and Evolutionary Origins of Color Diversity in Great
Star Coral Montastraea cavernosa (Scleractinia: Faviida).";
RL Mol. Biol. Evol. 20:1125-1133 (2003).
DR EMBL; AY181556; AAO61602.1; -.
SQ SEQUENCE 225 AA; 25843 MW; 13708587B7D93E35 CRC64;
```

```
Query Match 53.0%; Score 672; DB 5; Length 225;
Best Local Similarity 53.4%; Pred. No. 9.7e-56;
Matches 117; Conservative 45; Mismatches 57; Indels 0; Gaps 0;

Qy 1 SVIAKQMTYKVMGTVNGHYFEVGEKGKPYGEQTVRLAVTKGGLPFPAMDILSPQC 60
Db 2 SVIKSVMKIKLHMDDGVNGHKFMTITGEGEGKPFEGTHTIILKVKEGGGLPFAYDILTAA 61
Qy 61 QYGSIPFTKYPEDIPDYVKQSPGRTYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKS 120
Db 62 QYGNRVFTKYPDKIDPDKFQSPGYSWERSMTFEDQGVCTVSDIKLEGGDCFFYEIRFY 121
Qy 121 GLNFPNGPVQKKTQGGNEPNTERLFARDGMLIGNNFMALKLEGGHYLCFFKSTYKARK 180
Db 122 GVNFPSSGPVMQKTKLWEPSTENMYRVDGVLKGDVNSRTLLLEGGHKHRCNFRSTYKAKK 181
Qy 181 PVKMPGYHYVDRKLDVTNNHNDYTSVEQREISIA 215
Db 182 GVLLPDEYHFVDRHRIELSHDNDYNTVRYENAVARPSML 220
```

```
RESULT 9
Q9U6Y3 PRELIMINARY; PRT; 266 AA.
AC Q9U6Y3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Fluorescent protein FP484.
OS Clavularia sp.
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Alcyonaria; Alcyonacea;
OC Clavulariidae; Clavularia.
OX NCBI_TaxID=86521;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=99436614; PubMed=10504696;
RA Matz M.V., Fradkov A.F., Labas Y.A., Savitsky A.P., Zaraisky A.G.,
RA Markelov M.L., Lukyanov S.A.;
RT "Fluorescent proteins from nonbioluminescent Anthozoa species.";
RL Nat. Biotechnol. 17:969-973 (1999).
DR EMBL; AF168424; AAF03374.1; -.
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP-like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PD01229; GFP.
DR ProDom; PD013756; Green fl protein; 1.
SQ SEQUENCE 266 AA; 30450 MW; B4E97406B2708854 CRC64;
```

```
Query Match 52.8%; Score 670; DB 5; Length 266;
Best Local Similarity 54.6%; Pred. No. 1.8e-55;
Matches 119; Conservative 39; Mismatches 60; Indels 0; Gaps 0;

Qy 2 VIAKQMTYKVMGTVNGHYFEVGEKGKPYGEQTVRLAVTKGGLPFPAMDILSPQC 61
Db 45 VIKPDMKIKLRMEGAVNGHNFVIEGEGKPYDGTHTINLVKEGAPLPFSYDILSNAFQ 104
Qy 62 YGSIPFTKYPEDIPDYVKQSPGRTYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHVXFSG 121
Db 105 YGNRALTKYPDDIADYFKQSPGYSWERTWTTFEDKGIVKVKSDISMEEDSFYIEIRFDG 164
Qy 122 LNEPFPNGPVQKKTQGGNEPNTERLFARDGMLIGNNFMALKLEGGHYLCFFKSTYKARKP 181
Db 165 MNFPNGPVQKKTQGGNEPNTERLFARDGMLIGNNFMALKLEGGHYLCFFKSTYKARKV 224
Qy 182 VKMPGYHYVDRKLDVTNNHNDYTSVEQREISIA 219
Db 225 VKLPDYHFVDRHRIELSHDNDYNTVRYENAVARYSLL 262
```

```
RESULT 10
Q7ZOW6 PRELIMINARY; PRT; 227 AA.
AC Q7ZOW6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Green fluorescent protein.
OS Montastraea cavernosa (great star coral).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Faviina; Faviidae; Montastraea.
OX NCBI_TaxID=63558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mc4;
RX MEDLINE=22689801; PubMed=12777529;
RA Kelmanson I.V., Matz M.V.;
RT "Molecular Basis and Evolutionary Origins of Color Diversity in Great
Star Coral Montastraea cavernosa (Scleractinia: Faviida).";
RL Mol. Biol. Evol. 20:1125-1133 (2003).
DR EMBL; AY181555; AAO61601.1; -.
SQ SEQUENCE 227 AA; 26055 MW; 4BE2CB64FDB0E890 CRC64;
```

```
Query Match 52.6%; Score 667.5; DB 5; Length 227;
Best Local Similarity 53.4%; Pred. No. 2.6e-55;
Matches 118; Conservative 43; Mismatches 57; Indels 3; Gaps 1;

Qy 1 SVIAKQMTYKVMGTVNGHYFEVGEKGKPYGEQTVRLAVTKGGLPFPAMDILSPQC 60
Db 2 SVIKPDMKIKLRMEGAVNGHNFVIEGEGKPYDGTHTINLVKEGAPLPFSYDILSNAFQ 104
Qy 61 QYGSIPFTKYPEDIPDYVKQSPGRTYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHV 117
Db 62 DYGNRVFAKYPDKIDPDKFQSPGYSWERSMTYEDQGIATNDITMKGVDVDCFLYKI 121
Qy 118 KFSGLNFPNGPVQKKTQGGNEPNTERLFARDGMLIGNNFMALKLEGGHYLCFFKSTYK 177
```

[illegible]

```

DR Pfam; PF01353; GFP; 1.
DR ProDom; PD013756; Green_fl_protein; 1.
SQ SEQUENCE 227 AA; 26017 MW; 5E312C54EA47F589 CRC64;

Query Match 52.5%; Score 665.5; DB 5; Length 227;
Best Local Similarity 53.4%; Pred. No. 4e-55;
Matches 118; Conservative 42; Mismatches 58; Indels 3; Gaps 1;

Qy 1 SVIAKQMTYKYVMSTGVNNGHYFEVGGDGKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2 SVIKPDMKIKLRMEGAVNGHKFVIEGGDGKPFEGTQSMDLTVKEGAPLPFPAYDILTTVF 61

Qy 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNPFEDGAVCTVSDSSIQ---GNCFIYHV 117
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
62 DYGNRVFAKYPQDIPDYFKQTFPEGYSWERSMTYEDQGICVATNDITLMKGVDVDFVYKI 121

Qy 118 KFSGLNFPNGPVMMKKTQGWEPNTERLFPARDGMLIGNFMALKEGGGHYLCBFKSTYK 177
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 RFDGVNFPANGPVMMKKTKEWSTERMYVVDGVLKGDVNMALLEGGHYRCDFKTYK 181

Qy 178 ARKPVKMPGYHYVDRLDVTNHNKDYTSVEQREISIAARKPL 218
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
182 AKKFVQLPDYHFVDHRIEILSHDKDYNKVKLYEHAHSGL 222

RESULT 13
Q7Z0W8 PRELIMINARY; PRT; 227 AA.
ID Q7Z0W8
AC Q7Z0W8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Green fluorescent protein.
OS Montastraea cavernosa (great star coral).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Faviina; Faviidae; Montastraea.
OX NCBI_taxid=63558;
RN RP
RP SEQUENCE FROM N.A.
RC STRAIN=mc2;
RX MEDLINE=22689801; PubMed=12777529;
RA Kelmanson I.V.; Matz M.V.;
RT "Molecular Basis and Evolutionary Origins of Color Diversity in Great
RT Star Coral Montastraea cavernosa (Scleractinia: Faviida).";
RL Mol. Biol. Evol. 20:1125-1133(2003).
DR EMBL; AY181553; AAO61599.1; -.
SQ SEQUENCE 227 AA; 26017 MW; 5E312C54EA47F589 CRC64;

Query Match 52.5%; Score 665.5; DB 5; Length 227;
Best Local Similarity 53.4%; Pred. No. 4e-55;
Matches 118; Conservative 42; Mismatches 58; Indels 3; Gaps 1;

Qy 1 SVIAKQMTYKYVMSTGVNNGHYFEVGGDGKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2 SVIKPDMKIKLRMEGAVNGHKFVIEGGDGKPFEGTQSMDLTVKEGAPLPFPAYDILTTVF 61

Qy 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNPFEDGAVCTVSDSSIQ---GNCFIYHV 117
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
62 DYGNRVFAKYPQDIPDYFKQTFPEGYSWERSMTYEDQGICVATNDITLMKGVDVDFVYKI 121

Qy 118 KFSGLNFPNGPVMMKKTQGWEPNTERLFPARDGMLIGNFMALKEGGGHYLCBFKSTYK 177
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 RFDGVNFPANGPVMMKKTKEWSTERMYVVDGVLKGDVNMALLEGGHYRCDFKTYK 181

Qy 178 ARKPVKMPGYHYVDRLDVTNHNKDYTSVEQREISIAARKPL 218
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
182 AKKFVQLPDYHFVDHRIEILSHDKDYNKVKLYEHAHSGL 222

RESULT 14
Q95UA7 PRELIMINARY; PRT; 225 AA.
AC Q95UA7
ID Q95UA7;

```

01-DEC-2001 (Tremblrel. 19, Created)
01-DEC-2001 (Tremblrel. 19, Last sequence update)
01-OCT-2003 (Tremblrel. 25, Last annotation update)
Cyan fluorescent protein (Fragment).
Montastraea cavernosa (great star coral),
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Faviina; Faviidae; Montastrea.
NCBI_TaxID=63558;
[1]
RN SEQUENCE FROM N.A.
RP Falkowski P.G., Sun Y.;
RA "Montastraea cavernosa fluorescent protein.";
RT Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
EL EMBL; AY056460; AALI7905.1; -
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP like.
DR InterPro; IPR000786; Green_fl_protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFLUORESCENT.
DR ProDom; PD013756; Green fl protein; 1.
FT NON TER 225 225
SQ SEQUENCE 225 AA; 25775 MW; 52DE2F716D083524 CRC64;

Query Match 52.4%; Score 664; DB 5; Length 225;
Best Local Similarity 53.4%; Pred. No. 5.6e-55;
Matches 117; Conservative 43; Mismatches 59; Indels 0; Gaps 0;

QY 1 SVIAKQMITYKYVMGTVNGHYVEFGDGKGPKYEGETVRLAVTKGGPLPEAWDILSPQC 60
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 2 SVIKSVWKIKLRMDGIYNGHKFMITGBEGGFPGTGTHIIUKVKEGGLPPAYDILTTF 61
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 61 QYGSIPTKYPEDIPDYVKOSFPGRYTWERIMNPFEDGAVCTVSNDSTQGNCFTYHVKFS 120
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 62 QYGNRVETYPKDIPDYFKOSFPBGYSWRSMTEDQGVCTVTSDIKLEGDCFFYRIFY 121
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 121 GLAPPPNGPVMKKTQTGWENTELRFARDGMIGNNFALKEGGHYLCFEKSTYKARK 180
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 122 GVNPSPGPVMKKTLEWEPSTENMYVRDGVLLGDVSRITLLEGDKHRCNFRSTYGAKK 181
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
QY 181 PVKMGPYHYVDKLDVTNNHNKDYTSVEQREISIAKPIIV 219
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :
Db 182 GVVLPEYHFVDHRIEILSHDKDVNTVEYNVARSML 220
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | :

RESULT 15
Q8T5F2 PRELIMINARY; PRT; 234 AA.
AC Q8T5F2;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
MCavFP 6.
OS Montastraea cavernosa (great star coral).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Scleractinia;
OC Faviina; Faviidae; Montastrea.
NCBI_TaxID=63558;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=21927629; PubMed=11929996;
RA Labas Y.A., Gurskaya N.G., Yanushevich Y.G., Pradkov A.F.,
LUkyanov K.A., Lukyanov S.A., Matz M.V.;
RT "Diversity and evolution of the green fluorescent protein family";
Proc. Natl. Acad. Sci. U.S.A. 99:4256-4261(2002).
DR EMBL; AV037769; AAK71335.1; -
DR GO; GO:0006091; P:energy pathways; IEA.
DR InterPro; IPR009017; GFP_like.
DR InterPro; IPR000786; Green fl protein.
DR Pfam; PF01353; GFP; 1.
DR PRINTS; PR01229; GFLUORESCENT.
DR ProDom; PD013756; Green fl protein; 1.
SQ SEQUENCE 234 AA; 26743 MW; C62C39EC8063A24D CRC64;

Query Match 51.7%; Score 655.5; DB 5; Length 234;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 06:12:47 ; Search time 87.5717 Seconds
(without alignments)
745.314 Million cell updates/sec

Title: US-09-890-463-3

Perfect score: 1268

Sequence: 1 SVIAKQMTYKYVMSGTVNGH.....SIARKPLVACCFRVRKSRHK 231

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 29Jan04: *
1: geneseqp1980s: *
2: geneseqp1990s: *
3: geneseqp2000s: *
4: geneseqp2001s: *
5: geneseqp2002s: *
6: geneseqp2003as: *
7: geneseqp2003bs: *
8: geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1268	100.0	231	3	Aay97149 Pigment p
2	1268	100.0	231	5	Abp70025 Colour Fa
3	1242	97.9	235	3	Aay97150 Pigment p
4	1242	97.9	235	5	Abp70042 Colour Fa
5	1212	95.6	235	5	Abp70026 Colour Fa
6	1202	94.8	220	5	Abp70037 Colour Fa
7	1184	93.4	226	5	Abp70036 Colour Fa
8	1170	92.3	220	5	Abp69941 Colour Fa
9	1170	92.3	220	5	Abp69940 Colour Fa
10	1167	92.0	220	5	Abp69939 Colour Fa
11	1164	91.8	220	5	Abp69925 Colour Fa
12	1162	91.6	220	5	Abp69930 Colour Fa
13	1162	91.6	223	5	Abp70032 Colour Fa
14	1156	91.2	220	5	Abp69926 Colour Fa
15	1156	91.2	223	5	Abp70030 Colour Fa
16	1154	91.0	220	5	Abp69956 Colour Fa
17	1153	90.9	220	5	Abp69928 Colour Fa
18	1151	90.8	221	5	Abp69992 Colour Fa
19	1151	90.8	221	5	Abp69991 Colour Fa
20	1151	90.8	223	5	Abp70029 Colour Fa
21	1149.5	90.7	222	5	Abp70028 Colour Fa
22	1148	90.5	220	5	Abp70007 Colour Fa
23	1147	90.5	221	5	Abp69967 Colour Fa
24	1147	90.5	221	5	Abp69966 Colour Fa
25	1147	90.5	221	5	Abp70004 Colour Fa

26	1147	90.5	223	5	Abp70033 Colour Fa
27	1147	90.5	235	5	Abp69963 Colour Fa
28	1147	90.5	235	5	Abp69961 Colour Fa
29	1146	90.4	221	5	Abp69978 Colour Fa
30	1143	90.1	220	5	Abp69952 Colour Fa
31	1143	90.1	220	5	Abp69959 Colour Fa
32	1143	90.1	220	5	Abp69934 Colour Fa
33	1143	90.1	220	5	Abp69958 Colour Fa
34	1143	90.1	220	5	Abp69937 Colour Fa
35	1143	90.1	220	5	Abp69935 Colour Fa
36	1143	90.1	221	5	Abp69972 Colour Fa
37	1141.5	90.0	222	5	Abp70027 Colour Fa
38	1140	89.9	220	5	Abp69965 Colour Fa
39	1140	89.9	221	5	Abp70002 Colour Fa
40	1139	89.8	220	5	Abp69938 Colour Fa
41	1139	89.8	223	5	Abp70031 Colour Fa
42	1138	89.7	220	5	Abp69964 Colour Fa
43	1137	89.7	220	5	Abp69993 Colour Fa
44	1137	89.7	220	5	Abp69989 Colour Fa
45	1137	89.7	221	5	Abp69986 Colour Fa

ALIGNMENTS

RESULT 1

AAY97149

ID AAY97149 standard; protein; 231 AA.

XX AC

AAY97149;

XX AC

DT 04-DEC-2000 (first entry)

XX AC

DE Pigment protein from coral tissue POC3.

XX DE

KW N-terminal; pigment protein from coral tissue; POC3; fluorescence;

KW tissue marker; fluorescent marker; dyestuff; sunscreen; ultra violet;

KW UV filter; POC3.

XX KW

OS Acropora aspera.

XX OS

FH Key Location/Qualifiers

FT Misc-difference 61. .63

FT /label= Chromophore_motif

FT FT

FT Misc-difference 158 /note= "critical residue in the vicinity of the fluorophore"

FT FT

FT Misc-difference 192 /note= "critical residue in the vicinity of the fluorophore"

FT FT

FT Misc-difference 210 /note= "critical residue in the vicinity of the fluorophore"

FT FT

FT Misc-difference 210 /note= "critical residue in the vicinity of the fluorophore"

FT FT

PN WO200046233-A1.

XX PN

XX PD

PD 10-AUG-2000.

XX PD

PF 02-FEB-2000; 2000WO-AU0000056.

XX PF

XX PR

PR 02-FEB-1999; 99AU-000008463.

XX PR

XX PA

PA (UNSY) UNIV SYDNEY.

XX PA

XX PI

PI Hoegh-Guldberg O, Dove S;

XX PI

XX DR

DR WPI: 2000-532892/48.

XX DR

XX XX

XX PT

PT Novel pigment protein derived from corals capable of emitting fluorescence upon irradiation by incident light useful as tissue marker, fluorescent marker or general dyestuff.

XX PT

PS Claim 13; Page 42-43; 49pp; English.

XX cDNA libraries were constructed from a blue pigmented coral, *Acropora*

CC aspera to isolate sequences encoding polypeptides with N-terminal

CC sequences as shown in AAY97147-48. Pigment protein from coral tissue

CC (PPCT) is capable of emitting fluorescence upon irradiation by incident

CC light whose maximal absorbance is in the range of 320-600 nm and a

CC maximal fluorescence emission is in the range of 300-700 nm. PPCT may be

CC used as a tissue marker, fluorescent marker (e.g. to follow gene

CC expression in transformed tissues) or general dyestuff (all claimed).

CC PPCT may also be used in sunscreen formulations or UV filters (both

CC claimed)

XX Sequence 231 AA;

SQ Query Match 100.0%; Score 1268; DB 3; Length 231;

Best Local Similarity 100.0%; Pred. No. 1.3e-128;

Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVIAKQMTYKYVMSTGVNGHYFEVGDGKPKYEGEQTVRLAVTKGGPLPFAWDILSPQC 60

DB 1 SVIAKQMTYKYVMSTGVNGHYFEVGDGKPKYEGEQTVRLAVTKGGPLPFAWDILSPQC 60

QY 61 QYGSIPFTKYPEDIPDYVQKSPFGRYTWERIMNPDGAVCTVSDSSIQGNCFIYHVKFS 120

DB 61 QYGSIPFTKYPEDIPDYVQKSPFGRYTWERIMNPDGAVCTVSDSSIQGNCFIYHVKFS 120

QY 121 GLNFPNGPVVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGCHYLCEFKSYKARK 180

DB 121 GLNFPNGPVVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGCHYLCEFKSYKARK 180

QY 181 PVKMPGVHYVDRKLDVTNNKDYTSVEQREISIAKPKLVACCFPRVKSRRHK 231

DB 181 PVKMPGVHYVDRKLDVTNNKDYTSVEQREISIAKPKLVACCFPRVKSRRHK 231

RESULT 2

ID ABP70025 standard; protein; 231 AA.

XX AC ABP70025;

XX DT 06-AUG-2003 (revised)

XX DT 22-JAN-2003 (first entry)

XX DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 201.

XX KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;

XX KW chromophore; biomatrix; transgenic animal; colouring agent;

XX KW flower industry; expression marker; reporter molecule; photon trap;

XX KW UV sink; sunscreen.

XX OS Unidentified.

XX WO200270703-A2.

XX PD 12-SEP-2002.

XX PF 01-MAR-2002; 2002WO-GB000928.

XX PR 02-MAR-2001; 2001US-0273227P.

XX PR 21-MAR-2001; 2001AU-00003874.

XX PR 15-OCT-2001; 2001US-0329816P.

XX PA (NUFA-) NUFARM LTD.

XX PA (UYOU) UNIV QUEENSLAND.

XX PA (JONE/) JONES E L.

XX FI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;

XX FI Hoegh-Guldberg IO, Prescott M;

XX DR WPI; 2002-740765/80.

PT Novel color-facilitating molecule for producing a biomatrix, has a

PT polypeptide which alone/along with molecules imparts altered visual

PT characteristics to cells in the absence of excitation by extraneous non-

XX white light.

XX Claim 6; Page 478; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)

CC comprising a polypeptide which, in a cell, alone or together with one or

CC more other molecules imparts an altered visual characteristic to the cell

CC when visualised by a human eye in the absence of excitation by extraneous

CC non-white light or particle emission. CFMs are useful for producing a

CC transgenic animal which exhibits a novel colour e.g. sheep with blue or

CC red coloured fleece. They are useful for producing coloured plant

CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other

CC uses include transducing or intensifying an image, providing additional

CC light for growing phototropic organisms e.g. algae and/or corals, for

CC coating materials that experience UV damage e.g. plastics and car

CC upholstery. CFMs are useful in the flower industry, in the development of

CC new varieties of flowering plants. Other contemplated uses include,

CC expression markers, general reporter molecules, photon traps, UV sinks or

CC in sunscreens. CFMs modify visible colour in edible and/or ornamental

CC fungal species, and in fruits and vegetables to enhance their

CC marketability. CFMs embedded in a gel matrix improve image quality in

CC situations of distorted light spectra (biomatrix). The first all-protein

CC chromophore to be isolated was Green Fluorescent protein (GFP). The

CC sequences given in records ABP69924-ABP70048 represent CFM related amino

CC acid sequences. (Updated on 06-AUG-2003 to correct OS field.)

XX Sequence 231 AA;

SQ Query Match 100.0%; Score 1268; DB 5; Length 231;

Best Local Similarity 100.0%; Pred. No. 1.3e-128;

Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVIAKQMTYKYVMSTGVNGHYFEVGDGKPKYEGEQTVRLAVTKGGPLPFAWDILSPQC 60

DB 1 SVIAKQMTYKYVMSTGVNGHYFEVGDGKPKYEGEQTVRLAVTKGGPLPFAWDILSPQC 60

QY 61 QYGSIPFTKYPEDIPDYVQKSPFGRYTWERIMNPDGAVCTVSDSSIQGNCFIYHVKFS 120

DB 61 QYGSIPFTKYPEDIPDYVQKSPFGRYTWERIMNPDGAVCTVSDSSIQGNCFIYHVKFS 120

QY 121 GLNFPNGPVVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGCHYLCEFKSYKARK 180

DB 121 GLNFPNGPVVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGCHYLCEFKSYKARK 180

QY 181 PVKMPGVHYVDRKLDVTNNKDYTSVEQREISIAKPKLVACCFPRVKSRRHK 231

DB 181 PVKMPGVHYVDRKLDVTNNKDYTSVEQREISIAKPKLVACCFPRVKSRRHK 231

RESULT 3

AAAY97150

ID AAY97150 standard; protein; 235 AA.

XX AC AAY97150;

XX DT 04-DEC-2000 (first entry)

XX DE Pigment protein from coral tissue POC4.

XX KW N-terminal; pigment protein from coral tissue; PPCT; fluorescence;

XX KW tissue marker; fluorescent marker; dyestuff; sunscreen; ultra violet;

XX KW UV filter; POC3.

XX OS *Acropora aspera*.

XX FH Key Location/Qualifiers

XX FT Misc-difference 61..63

XX FT Misc-difference /label= Chromophore_motif

XX FT Misc-difference 158

XX FT /note= "critical residue in the vicinity of the

FT fluorophore"
 FT Misc-difference 192
 FT /note= "critical residue in the vicinity of the
 FT fluorophore"
 FT Misc-difference 210
 FT /note= "critical residue in the vicinity of the
 FT fluorophore"
 XX WO200046233-A1.
 XX 10-AUG-2000.
 XX 02-FEB-2000; 2000WO-AU000056.
 XX 02-FEB-1999; 99AU-00008463.
 XX (UNSY) UNIV SYDNEY.
 XX Hoegh-Guldberg O, Dove S;
 PI MPI; 2000-532892/48.
 DR N-PSDB; AAA52083.
 XX Novel pigment protein derived from corals capable of emitting
 PT fluorescence upon irradiation by incident light useful as tissue marker,
 PT fluorescent marker or general dyestuff.
 XX Claim 13; Page 43-44; 49pp; English.
 XX cDNA libraries were constructed from a blue pigmented coral, Acropora
 CC aspera to isolate sequences encoding polypeptides with N-terminal
 CC sequences as shown in AA97147-48. Pigment protein from coral tissue
 CC (PPCT) is capable of emitting fluorescence upon irradiation by incident
 CC light whose maximal absorbance is in the range of 320-600 nm and a
 CC maximal fluorescence emission is in the range of 300-700 nm. PPCT may be
 CC used as a tissue marker, fluorescent marker (e.g. to follow gene
 CC expression in transformed tissues) or general dyestuff (all claimed).
 CC PPCT may also be used in sunscreen formulations or UV filters (both
 CC claimed)
 XX SQ Sequence 235 AA;
 Query Match 97.9%; Score 1242; DB 3; Length 235;
 Best Local Similarity 98.3%; Pred. No. 8.9e-126;
 Matches 227; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SVIAKQMTYKVMGTVNGHYFEVGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
 Db 1 SVIAKQMTYKVMGTVNGHYFEVGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
 QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFEDGAVCTVNSDSSIQGNCFIYHVKFS 120
 Db 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFEDGAVCTVNSDSSIQGNCFIYHVKFS 120
 QY 121 GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNFNFMALKEGGHYLCBPKSTYKARK 180
 Db 121 GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNFNFMALKEGGHYLCBPKSTYKARK 180
 QY 181 PVKMPGHHYVDRKLDVTNHNKDYTSVEQREISIAKPLVACCFRVRKSHK 231
 Db 181 PVKMPGHHYVDRKLDVTNHNKDYTSVEQREISIAKPLVACCFRVRKSHK 231
 RESULT 4
 ABP70042
 ID ABP70042 standard; protein; 235 AA.
 XX AC ABP70042;
 XX 22-JAN-2003 (first entry)
 DT Colour Facilitating molecule (CFM) related sequence #SEQ ID 245.
 DE
 XX

KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunscreen.
 XX Acropora aspera.
 XX WO200270703-A2.
 XX 12-SEP-2002.
 XX 01-MAR-2002; 2002WO-GB000928.
 XX 02-MAR-2001; 2001US-0273227P.
 XX 21-MAR-2001; 2001AU-00003874.
 XX 15-OCT-2001; 2001AU-0329816P.
 XX (NUFA-) NUFARM LTD.
 XX (UYOU) UNIV QUEENSLAND.
 XX (JONE/) JONES E L.
 XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
 PI Hoegh-Guldberg IO, Prescott M;
 XX MPI; 2002-740765/80.
 XX Novel color-facilitating molecule for producing a biomatrix, has a
 PT polypeptide which alone/along with molecules imparts altered visual
 PT characteristics to cells in the absence of excitation by extraneous non-
 PT white light.
 XX Example 20; Page 502-503; 510pp; English.
 XX The invention relates to an isolated colour-facilitating molecule (CFM)
 CC comprising a polypeptide which, in a cell, alone or together with one or
 CC more other molecules imparts an altered visual characteristic to the cell
 CC when visualised by a human eye in the absence of excitation by extraneous
 CC non-white light or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant
 CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development of
 CC new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC sequences given in records ABP69924-ABP70048 represent CFM related amino
 CC acid sequences
 XX SQ Sequence 235 AA;
 Query Match 97.9%; Score 1242; DB 5; Length 235;
 Best Local Similarity 98.3%; Pred. No. 8.9e-126;
 Matches 227; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 1 SVIAKQMTYKVMGTVNGHYFEVGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
 Db 1 SVIAKQMTYKVMGTVNGHYFEVGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
 QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFEDGAVCTVNSDSSIQGNCFIYHVKFS 120
 Db 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFEDGAVCTVNSDSSIQGNCFIYHVKFS 120
 QY 121 GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNFNFMALKEGGHYLCBPKSTYKARK 180
 Db 121 GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNFNFMALKEGGHYLCBPKSTYKARK 180

CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development of
 CC new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC sequences given in records ABP69924-ABP70048 represent CFM related amino
 CC acid sequences. (Updated on 06-AUG-2003 to correct OS field.)
 XX Sequence 220 AA;
 SQ

Query Match 94.8%; Score 1202; DB 5; Length 220;
 Best Local Similarity 100.0%; Pred. No. 1.7e-121;
 Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVIAKQMTYKVMSTVNGHYFEVGDGKPGYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
 Db 1 SVIAKQMTYKVMSTVNGHYFEVGDGKPGYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
 Qy 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFDGAVCTVSDSSIQGNCFIYHVKFS 120
 Db 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFDGAVCTVSDSSIQGNCFIYHVKFS 120
 Qy 121 GLNFPNGPVVMQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCEPKSTYKARK 180
 Db 121 GLNFPNGPVVMQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCEPKSTYKARK 180
 Qy 181 PVKMPGYHYVDKLDVTNHNKDYTSVEQREISIAARKPLVA 220
 Db 181 PVKMPGYHYVDKLDVTNHNKDYTSVEQREISIAARKPLVA 220

RESULT 7
 ABP70036
 ID ABP70036 standard; protein; 226 AA.
 AC
 XX
 DT 06-AUG-2003 (revised)
 DT 22-JAN-2003 (first entry)
 XX
 DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 238.
 XX
 KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunsreen.
 XX
 OS Unidentified.
 XX
 XX
 XX WO200270703-A2.
 XX
 XX 12-SEP-2002.
 XX
 XX 01-MAR-2002; 2002WO-GB000928.
 XX
 XX 02-MAR-2001; 2001US-0273227P.
 XX 21-MAR-2001; 2001AU-00003874.
 XX 15-OCT-2001; 2001US-0329816P.
 XX
 XX (NUFA-) NUFARM LTD.
 XX (YOU) UNIV QUEENSLAND.
 XX (JONE/) JONES E L.
 XX
 XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
 XX PI Hoegh-Guldberg IO, Prescott M;
 XX WPI; 2002-740765/80.
 DR

XX Novel color-facilitating molecule for producing a biomatrix, has a
 PT polypeptide which alone/along with molecules imparts altered visual
 PT characteristics to cells in the absence of excitation by extraneous non-
 PT white light.
 XX Example 19; Page 496-497; 510pp; English.
 XX
 CC The invention relates to an isolated colour-facilitating molecule (CFM)
 CC comprising a polypeptide which, in a cell, alone or together with one or
 CC more other molecules imparts an altered visual characteristic to the cell
 CC when visualised by a human eye in the absence of excitation by extraneous
 CC non-white light or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant
 CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development of
 CC new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC sequences given in records ABP69924-ABP70048 represent CFM related amino
 CC acid sequences. (Updated on 06-AUG-2003 to correct OS field.)
 XX Sequence 226 AA;
 SQ

Query Match 93.4%; Score 1184; DB 5; Length 226;
 Best Local Similarity 96.0%; Pred. No. 1.6e-113;
 Matches 217; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 1 SVIAKQMTYKVMSTVNGHYFEVGDGKPGYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
 Db 1 SVIAKQMTYKVMSTVNGHYFEVGDGKPGYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
 Qy 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFDGAVCTVSDSSIQGNCFIYHVKFS 120
 Db 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFDGAVCTVSDSSIQGNCFIYHVKFS 120
 Qy 121 GLNFPNGPVVMQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCEPKSTYKARK 180
 Db 121 GLNFPNGPVVMQKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCEPKSTYKARK 180
 Qy 181 PVKMPGYHYVDKLDVTNHNKDYTSVEQREISIAARKPLVACCFRV 226
 Db 181 PVKMPGYHYVDKLDVTNHNKDYTSVEQREISIAARKPLVACCFRV 226

RESULT 8
 ABP69941
 ID ABP69941 standard; protein; 220 AA.
 XX
 XX AC ABP69941;
 XX
 XX 22-JAN-2003 (first entry)
 XX
 XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 54.
 XX
 XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunsreen.
 XX
 XX Millepora sp.
 XX
 XX WO200270703-A2.
 XX
 XX 12-SEP-2002.
 PD

```

XX PF 01-MAR-2002; 2002WO-GB0000928.
XX PR 02-MAR-2001; 2001US-0273227P.
XX PR 21-MAR-2001; 2001AU-00003874.
XX PR 15-OCT-2001; 2001US-0329816P.
XX PA (NUFA-) NUFARM LTD.
XX PA (UYQU ) UNIV QUEENSLAND.
XX PA (JONE/) JONES E L.
XX PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
XX PI Hoegh-Guldberg IO, Prescott M;
XX DR WPI; 2002-740765/80.
XX PF Novel color-facilitating molecule for producing a biomatrix, has a
XX PT polypeptide which alone/along with molecules imparts altered visual
XX PT characteristics to cells in the absence of excitation by extraneous non-
XX PT white light.
XX PS Claim 5; Page 330-331; 510pp; English.
XX CC The invention relates to an isolated colour-facilitating molecule (CFM)
XX CC comprising a polypeptide which, in a cell, alone or together with one or
XX CC more other molecules imparts an altered visual characteristic to the cell
XX CC when visualised by a human eye in the absence of excitation by extraneous
XX CC non-white light or particle emission. CFMs are useful for producing a
XX CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
XX CC red coloured fleece. They are useful for producing coloured plant
XX CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
XX CC uses include transducing or intensifying an image, providing additional
XX CC light for growing phototropic organisms e.g. algae and/or corals, for
XX CC coating materials that experience UV damage e.g. plastics and car
XX CC upholstery. CFMs are useful in the flower industry, in the development of
XX CC new varieties of flowering plants. Other contemplated uses include,
XX CC expression markers, general reporter molecules, photon traps, UV sinks or
XX CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
XX CC fungal species, and in fruits and vegetables to enhance their
XX CC marketability. CFMs embedded in a gel matrix improve image quality in
XX CC situations of distorted light spectra (biomatrix). The first all-protein
XX CC chromophore to be isolated was Green Fluorescent protein (GFP). The
XX CC sequences given in records ABP69924-ABP70048 represent CFM related amino
XX CC acid sequences
XX PS Sequence 220 AA;
XX CC
XX CC Query Match          92.3%; Score 1170; DB 5; Length 220;
XX CC Best Local Similarity 97.7%; Pred. No. 5.1e-118;
XX CC Matches 215; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX QY 1 SVIAKQMTYKYVMSGTVNGHYFEVEGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
XX DB 1 SVIAKQMTYKYVMSGTVNGHYFEVEGDGKGPYEGEQTVRLTVTKGGPLPFAWDILSPQS 60
XX QY 61 QYGSIPFTKYPEDIPDYVVKQSPGRTYWERIMNFEDGAVCTVSDSSIQGNCFTYHVKFS 120
XX DB 61 QYGSIPFTKYPEDIPDYVVKQSPGRTYWERIMNFEDGAVCTVSDSSIQGNCFTYHVKFS 120
XX QY 121 GLNFPFPGVPMQKKTQGWEPNTERLFARDGMLIGNFMAKLEGGGHVLCFFKSTYKARK 180
XX DB 121 GLNFPFPGVPMQKKTQGWEPNTERLFARDGMLIGNFMAKLEGGGHVLCFFKSTYKARK 180
XX QY 181 PVKMPGHVYVDRKLDVTNNHNDYTSVEQRETSIARKPLVA 220
XX DB 181 PVKMPGHVYVDRKLDVTNNHNDYTSVEQRETSIARKPLVA 220
XX RESULT 9
XX ABP69940
XX ID ABP69940 standard; protein; 220 AA.
XX AC ABP69940;

```

```

XX DT 22-JAN-2003 (first entry)
XX DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 52.
XX KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
XX KW chromophore; biomatrix; transgenic animal; colouring agent;
XX KW flower industry; expression marker; reporter molecule; photon trap;
XX KW UV sink; sunscreen.
XX OS Miliepora sp.
XX PN WO200270703-A2.
XX PD 12-SEP-2002.
XX PF 01-MAR-2002; 2002WO-GB0000928.
XX PR 02-MAR-2001; 2001US-0273227P.
XX PR 21-MAR-2001; 2001AU-00003874.
XX PR 15-OCT-2001; 2001US-0329816P.
XX PA (NUFA-) NUFARM LTD.
XX PA (UYQU ) UNIV QUEENSLAND.
XX PA (JONE/) JONES E L.
XX PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
XX PI Hoegh-Guldberg IO, Prescott M;
XX DR WPI; 2002-740765/80.
XX PF Novel color-facilitating molecule for producing a biomatrix, has a
XX PT polypeptide which alone/along with molecules imparts altered visual
XX PT characteristics to cells in the absence of excitation by extraneous non-
XX PT white light.
XX PS Claim 5; Page 327-328; 510pp; English.
XX CC The invention relates to an isolated colour-facilitating molecule (CFM)
XX CC comprising a polypeptide which, in a cell, alone or together with one or
XX CC more other molecules imparts an altered visual characteristic to the cell
XX CC when visualised by a human eye in the absence of excitation by extraneous
XX CC non-white light or particle emission. CFMs are useful for producing a
XX CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
XX CC red coloured fleece. They are useful for producing coloured plant
XX CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
XX CC uses include transducing or intensifying an image, providing additional
XX CC light for growing phototropic organisms e.g. algae and/or corals, for
XX CC coating materials that experience UV damage e.g. plastics and car
XX CC upholstery. CFMs are useful in the flower industry, in the development of
XX CC new varieties of flowering plants. Other contemplated uses include,
XX CC expression markers, general reporter molecules, photon traps, UV sinks or
XX CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
XX CC fungal species, and in fruits and vegetables to enhance their
XX CC marketability. CFMs embedded in a gel matrix improve image quality in
XX CC situations of distorted light spectra (biomatrix). The first all-protein
XX CC chromophore to be isolated was Green Fluorescent protein (GFP). The
XX CC sequences given in records ABP69924-ABP70048 represent CFM related amino
XX CC acid sequences
XX PS Sequence 220 AA;
XX CC
XX CC Query Match          92.3%; Score 1170; DB 5; Length 220;
XX CC Best Local Similarity 97.7%; Pred. No. 5.1e-118;
XX CC Matches 215; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
XX QY 1 SVIAKQMTYKYVMSGTVNGHYFEVEGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
XX DB 1 SVIAKQMTYKYVMSGTVNGHYFEVEGDGKGPYEGEQTVRLTVTKGGPLPFAWDILSPQS 60
XX QY 61 QYGSIPFTKYPEDIPDYVVKQSPGRTYWERIMNFEDGAVCTVSDSSIQGNCFTYHVKFS 120
XX DB 61 QYGSIPFTKYPEDIPDYVVKQSPGRTYWERIMNFEDGAVCTVSDSSIQGNCFTYHVKFS 120

```

QY 121 GLNPPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGHYLCEFKSTYKARK 180
Db 121 GLNPPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGHYLCEFKSTYKARK 180
QY 181 PVKMPGHHYVDRKLDVTHNHNKDYTSVQRELSIARKPLVA 220
Db 181 PVKMPGHHYVDRKLDVTHNHNKDYTSVQRELSIARKPVVA 220

RESULT 10
ID ABP69939 standard; protein; 220 AA.
XX AC ABP69939;
XX DT 22-JAN-2003 (first entry)
XX DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 50.
XX KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
XX KW chromophore; biomatrix; transgenic animal; colouring agent;
XX KW flower industry; expression marker; reporter molecule; photon trap;
XX KW UV sink; sunscreen.
XX OS Millepora sp.
XX FN WO200270703-A2.
XX PD 12-SEP-2002.
XX PF 01-MAR-2002; 2002WO-GB0000928.
XX PR 02-MAR-2001; 2001US-0273227P.
XX PR 21-MAR-2001; 2001AU-00003874.
XX PR 15-OCT-2001; 2001US-0329816P.
XX PA (NUFA-) NUFARM LTD.
XX PA (UYQU) UNIV QUEENSLAND.
XX PA (JONE/) JONES E L.
XX PI Jones EL, Kavan M, Brugliera F, Mason J, Dove SG;
XX PI Hoegh-Guldberg IO, Prescott M;
XX DR WPI; 2002-740765/80.
XX PT Novel color-facilitating molecule for producing a biomatrix, has a
XX PT polypeptide which alone/along with molecules imparts altered visual
XX PT characteristics to cells in the absence of excitation by extraneous non-
XX PT white light.
XX PS Claim 5; Page 325-326; 510pp; English.
XX CC The invention relates to an isolated colour-facilitating molecule (CFM)
XX CC comprising a polypeptide which, in a cell, alone or together with one or
XX CC more other molecules imparts an altered visual characteristic to the cell
XX CC when visualised by a human eye in the absence of excitation by extraneous
XX CC non-white light or particle emission. CFMs are useful for producing a
XX CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
XX CC red coloured fleece. They are useful for producing coloured plant
XX CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
XX CC uses include transducing or intensifying an image, providing additional
XX CC light for growing phototropic organisms e.g. algae and/or corals, for
XX CC coating materials that experience UV damage e.g. plastics and car
XX CC upholstery. CFMs are useful in the flower industry, in the development of
XX CC new varieties of flowering plants. Other contemplated uses include,
XX CC expression markers, general reporter molecules, photon traps, UV sinks or
XX CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
XX CC fungal species, and in fruits and vegetables to enhance their
XX CC marketability. CFMs embedded in a gel matrix improve image quality in
XX CC situations of distorted light spectra (biomatrix). The first all-protein
XX CC chromophore to be isolated was Green Fluorescent protein (GFP). The
XX CC sequences given in records ABP69924-ABP70048 represent CFM related amino

CC acid sequences
XX Sequence 220 AA;
SQ Query Match 92.0%; Score 1167; DB 5; Length 220;
Best Local Similarity 97.3%; Pred. No. 1.1e-117; Indels 0; Gaps 0;
Matches 214; Conservative 2; Mismatches 4;
QY 1 SVIAKQMTYKYVMSGTVNGHYFEVGGKGPYEGEOTVRLAVTKGGPLPFAWDILSPQC 60
Db 1 SVIAKQMTYKYVMSGTVNGHYFEVGGKGPYEGEOTVRLAVTKGGPLPFAWDILSPQS 60
QY 61 QYGSIPFTKYPEDIPDYVKQSFPGRYTWERIMNPFEDGAVCTVSNDSIQGNCFIYHVKFS 120
Db 61 QYGSIPFTKYPEDIPDYVKQSFPGRYTWERIMNPFEDGAVCTVSNDSIQGNCFIYHVKFS 120
QY 121 GLNPPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGHYLCEFKSTYKARK 180
Db 121 GLNPPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGHYLCEFKSTYKARK 180
QY 181 PVKMPGHHYVDRKLDVTHNHNKDYTSVQRELSIARKPLVA 220
Db 181 PVKMPGHHYVDRKLDVTHNHNKDYTSVQRELSIARKPVVA 220

RESULT 11
ID ABP69925 standard; protein; 220 AA.
XX AC ABP69925;
XX DT 22-JAN-2003 (first entry)
XX DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 22.
XX KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
XX KW chromophore; biomatrix; transgenic animal; colouring agent;
XX KW flower industry; expression marker; reporter molecule; photon trap;
XX KW UV sink; sunscreen.
XX OS Acropora aspera.
XX FN WO200270703-A2.
XX PD 12-SEP-2002.
XX PF 01-MAR-2002; 2002WO-GB0000928.
XX PR 02-MAR-2001; 2001US-0273227P.
XX PR 21-MAR-2001; 2001AU-00003874.
XX PR 15-OCT-2001; 2001US-0329816P.
XX PA (NUFA-) NUFARM LTD.
XX PA (UYQU) UNIV QUEENSLAND.
XX PA (JONE/) JONES E L.
XX PI Jones EL, Kavan M, Brugliera F, Mason J, Dove SG;
XX PI Hoegh-Guldberg IO, Prescott M;
XX DR WPI; 2002-740765/80.
XX PT Novel color-facilitating molecule for producing a biomatrix, has a
XX PT polypeptide which alone/along with molecules imparts altered visual
XX PT characteristics to cells in the absence of excitation by extraneous non-
XX PT white light.
XX PS Claim 5; Page 286-287; 510pp; English.
XX CC The invention relates to an isolated colour-facilitating molecule (CFM)
XX CC comprising a polypeptide which, in a cell, alone or together with one or
XX CC more other molecules imparts an altered visual characteristic to the cell
XX CC when visualised by a human eye in the absence of excitation by extraneous
XX CC non-white light or particle emission. CFMs are useful for producing a

transgenic animal which exhibits a novel colour e.g. sheep with blue or red coloured fleece. They are useful for producing coloured plant extracts, e.g. flavouring, beverage or juice or colouring agent. Other uses include transducing or intensifying an image, providing additional light for growing phototropic organisms e.g. algae and/or corals, for coating materials that experience UV damage e.g. plastics and car upholstery. CFMs are useful in the flower industry, in the development of new varieties of flowering plants. Other contemplated uses include, expression markers, general reporter molecules, photon traps, UV sinks or in sunscreens. CFMs modify visible colour in edible and/or ornamental fungal species, and in fruits and vegetables to enhance their marketability. CFMs embedded in a gel matrix improve image quality in situations of distorted light spectra (biomatrix). The first all-protein chromophore to be isolated was Green Fluorescent protein (GFP). The sequences given in records ABP69924-ABP70048 represent CFM related amino acid sequences

XX
SQ Sequence 220 AA;

Query Match 91.8%; Score 1164; DB 5; Length 220;
Best Local Similarity 97.3%; Pred. No. 2.3e-117;
Matches 214; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 SVIAQMTYKVMSTVNGHYFEVGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
DB 1 SVIAQMTYKVMSTVNGHYFEVGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQS 60
QY 61 QYGSIPFTKYPEDIPDYVKQSPFGYTWERIMNFDGAVCTVSDSSIQGNCFIYHVKFS 120
DB 61 QYGSIPFTKYPEDIPDYVKQSPFGYTWERIMNFDGAVCTVSDSSIQGNCFIYHVKFS 120
QY 121 GLNPPNGPVVMQKKTQGWEPNTERLFARDGMLIGNFMAKLEGGHYLCBFKSTYKARK 180
DB 121 GLNPPNGPVVMQKKTQGWEPNTERLFARDGMLIGNFMAKLEGGHYLCBFKSTYKARK 180
QY 181 PVKMPGHHYVDRKLDVTNNHNDYTSVEQREISIAKPLVA 220
DB 181 PVKMPGHHYVDRKLDVTNNHNDYTSVEQREISIAKPLVA 220

RESULT 12
ID ABP69930 standard; protein; 220 AA.
AC ABP69930;
XX
XX
DT 06-AUG-2003 (revised)
XX
XX 22-JAN-2003 (first entry)
DE
DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 32.
XX
XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
KW UV sink; sunsreen.
XX
OS Caulastrea sp.
XX
XX WO200270703-A2.
XX
XX 12-SEP-2002.
XX
XX 01-MAR-2002; 2002WO-GB000928.
XX
XX 02-MAR-2001; 2001US-0273227P.
PR 21-MAR-2001; 2001AU-00003874.
PR 15-OCT-2001; 2001US-0329816P.
XX
XX (NUFA-) NUFARM LTD.
PA (UYQU) UNIV QUEENSLAND.
PA (JONE/) JONES E L.
XX
XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;

PI Hoegh-Guldberg IO, Prescott M;
XX
XX WPI; 2002-740765/80.

XX Novel color-facilitating molecule for producing a biomatrix, has a
PT polypeptide which alone/along with molecules imparts altered visual
PT characteristics to cells in the absence of excitation by extraneous non-
PT white light.

XX Claim 5; Page 298-299; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)
CC comprising a polypeptide which, in a cell, alone or together with one or
CC more other molecules imparts an altered visual characteristic to the cell
CC when visualised by a human eye in the absence of excitation by extraneous
CC non-white light or particle emission. CFMs are useful for producing a
CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
CC red coloured fleece. They are useful for producing coloured plant
CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
CC uses include transducing or intensifying an image, providing additional
CC light for growing phototropic organisms e.g. algae and/or corals, for
CC coating materials that experience UV damage e.g. plastics and car
CC upholstery. CFMs are useful in the flower industry, in the development of
CC new varieties of flowering plants. Other contemplated uses include,
CC expression markers, general reporter molecules, photon traps, UV sinks or
CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
CC fungal species, and in fruits and vegetables to enhance their
CC marketability. CFMs embedded in a gel matrix improve image quality in
CC situations of distorted light spectra (biomatrix). The first all-protein
CC chromophore to be isolated was Green Fluorescent protein (GFP). The
CC sequences given in records ABP69924-ABP70048 represent CFM related amino
CC acid sequences. (Updated on 06-AUG-2003 to correct OS field.)

XX Sequence 220 AA;

Query Match 91.6%; Score 1162; DB 5; Length 220;
Best Local Similarity 97.3%; Pred. No. 3.8e-117;
Matches 214; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 SVIAQMTYKVMSTVNGHYFEVGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
DB 1 SVIAQMTYKVMSTVNGHYFEVGDGKGPYEGEQTVRLAVTKGGPLPFAWDILSPQS 60
QY 61 QYGSIPFTKYPEDIPDYVKQSPFGYTWERIMNFDGAVCTVSDSSIQGNCFIYHVKFS 120
DB 61 QYGSIPFTKYPEDIPDYVKQSPFGYTWERIMNFDGAVCTVSDSSIQGNCFIYHVKFS 120
QY 121 GLNPPNGPVVMQKKTQGWEPNTERLFARDGMLIGNFMAKLEGGHYLCBFKSTYKARK 180
DB 121 GLNPPNGPVVMQKKTQGWEPNTERLFARDGMLIGNFMAKLEGGHYLCBFKSTYKARK 180
QY 181 PVKMPGHHYVDRKLDVTNNHNDYTSVEQREISIAKPLVA 220
DB 181 PVKMPGHHYVDRKLDVTNNHNDYTSVEQREISIAKPLVA 220

RESULT 13
ABP70032 standard; protein; 223 AA.
ID ABP70032
XX
XX 22-JAN-2003 (first entry)
AC ABP70032;
XX
XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 220.
XX
XX Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
KW UV sink; sunsreen.
XX
XX Simularia sp.

PN WO200270703-A2.
 XX 12-SEP-2002.
 XX 01-MAR-2002; 2002WO-GB000928.
 XX 02-MAR-2001; 2001US-0273227P.
 PR 21-MAR-2001; 2001AU-00003874.
 PR 15-OCT-2001; 2001US-0329816P.
 XX (NUFA-) NUFARM LTD.
 PA (OYQU) UNIV QUEENSLAND.
 PA (JONE/) JONES E L.
 XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
 PI Hoeigh-Guldberg IO, Prescott M;
 XX WPI; 2002-740765/80.
 XX Novel color-facilitating molecule for producing a biomatrix, has a
 PT polypeptide which alone/along with molecules imparts altered visual
 PT characteristics to cells in the absence of excitation by extraneous non-
 PT white light.
 XX Disclosure; Page 489; 510pp; English.
 XX The invention relates to an isolated colour-facilitating molecule (CFM)
 CC comprising a polypeptide which, in a cell, alone or together with one or
 CC more other molecules imparts an altered visual characteristic to the cell
 CC when visualised by a human eye in the absence of excitation by extraneous
 CC non-white light or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant
 CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development of
 CC new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC sequences given in records ABP69924-ABP70048 represent CFM related amino
 CC acid sequences
 XX Sequence 223 AA;
 SQ
 Query Match 91.6%; Score 1162; DB 5; Length 223;
 Best Local Similarity 97.3%; Pred. No. 3.8e-117; Indels 0; Gaps 0;
 Matches 214; Conservative 2; Mismatches 4;
 QY 1 SVIAKQMTYKVMSTGVNGHYFEVGDGKGKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
 Db 2 SVIAKQMTYKVMSTGVNGHYFEVGDGKGKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 61
 QY 61 QYGSIPFTKYLEDIPDYKQSPFGRYTWERINNFEDGAVCTVSDSSIQGNCFIYHVKFS 120
 Db 62 QYGSIPFTKYLEDIPDYKQSPFGRYTWERINNFEDGAVCTVSDSSIQGNCFIYHVKFS 121
 QY 121 GLNFPNGPVMOKKTQGMENPTEFLFARDGMLIGNFMALKEGGHYLCEFKSYKARK 180
 Db 122 GLNFPNGPVMOKKTQGMENPTEFLFARDGMLIGNFMALKEGGHYLCEFKSYKARK 181
 QY 181 PVKMPGYHYVDRKLDVTNNHNDYTSVEQREISIAKRLVA 220
 Db 182 PVKMPGYHYVDRKLDVTNNHNDYTSVEQREISIAKRLVA 221
 RESULT 14
 ABP69926

ID ABP69926 standard; protein; 220 AA.
 XX ABP69926;
 XX 22-JAN-2003 (first entry)
 XX Colour Facilitating molecule (CFM) related sequence #SEQ ID 24.
 DE Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 XX chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunsreen.
 XX Acropora aspera.
 XX WO200270703-A2.
 XX 12-SEP-2002.
 XX 01-MAR-2002; 2002WO-GB000928.
 PF 02-MAR-2001; 2001US-0273227P.
 XX 21-MAR-2001; 2001AU-00003874.
 PR 15-OCT-2001; 2001US-0329816P.
 XX (NUFA-) NUFARM LTD.
 PA (OYQU) UNIV QUEENSLAND.
 PA (JONE/) JONES E L.
 XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
 PI Hoeigh-Guldberg IO, Prescott M;
 XX WPI; 2002-740765/80.
 XX Novel color-facilitating molecule for producing a biomatrix, has a
 PT polypeptide which alone/along with molecules imparts altered visual
 PT characteristics to cells in the absence of excitation by extraneous non-
 PT white light.
 XX Claim 5; Page 289; 510pp; English.
 XX The invention relates to an isolated colour-facilitating molecule (CFM)
 CC comprising a polypeptide which, in a cell, alone or together with one or
 CC more other molecules imparts an altered visual characteristic to the cell
 CC when visualised by a human eye in the absence of excitation by extraneous
 CC non-white light or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant
 CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for
 CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development of
 CC new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC sequences given in records ABP69924-ABP70048 represent CFM related amino
 CC acid sequences
 XX Sequence 220 AA;
 SQ
 Query Match 91.2%; Score 1156; DB 5; Length 220;
 Best Local Similarity 96.4%; Pred. No. 1.7e-116;
 Matches 212; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 1 SVIAKQMTYKVMSTGVNGHYFEVGDGKGKPYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
 Db 1 SVIAKQMTYKVMSTGVNGHYFEVGDGKGKPYEGEQTVRLAVTKGGPLPFAWDILSPQS 60

QY 61 QYGSIPFTKYPEDIPDYVQKSFPGRYTWERIMNFEDGAVCTVSNDSISQGNCFIYHVKFS 120
Db |||||
QY 61 QYGSIPFTKYPEDIPDYVQKSFPGRYTWERIMNFEDGAVCTVSNDSISQGNCFIYHVKFS 120
Db |||||
QY 121 GLNFPNGPVNMQKTQGWEPNTERLFARDGMLIGNNFMALKEGGHYLCEFKSTYKARK 180
Db |||||
QY 181 PVKMPGHHYVDRKLDVTNNHNDYTSVEQREISIAKPLVA 220
Db |||||
QY 181 PVKMPGHHYVDRKLDVTNNHNDYTSVEQREISIAKPPVA 220

RESULT 15
ID ABP70030 standard; protein; 223 AA.
XX AC ABP70030;
XX DT 22-JAN-2003 (first entry)
XX DE Colour Facilitating molecule (CFM) related sequence #SEQ ID 216.
XX KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
KW UV sink; sunscreen.
XX OS Tubastrea sp.

XX WO200270703-A2.
XX PD 12-SEP-2002.
XX PF 01-MAR-2002; 2002WO-GB000928.
XX PR 02-MAR-2001; 2001US-0273227P.
PR 21-MAR-2001; 2001AU-0000387A.
PR 15-OCT-2001; 2001US-0329816P.
XX (NUFA-) NUFARM LTD.
PA (UYQU) UNIV QUEENSLAND.
PA (JONE/) JONES E L.
XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
PI Hoegh-Guldberg IO, Prescott M;
XX WPI; 2002-740765/80.
XX Novel color-facilitating molecule for producing a biomatrix, has a
PT polypeptide which alone/along with molecules imparts altered visual
PT characteristics to cells in the absence of excitation by extraneous non-
PT white light.

PS Example 18; Page 486; 510pp; English.
XX The invention relates to an isolated colour-facilitating molecule (CFM)
CC comprising a polypeptide which, in a cell, alone or together with one or
CC more other molecules imparts an altered visual characteristic to the cell
CC when visualised by a human eye in the absence of excitation by extraneous
CC non-white light or particle emission. CFMs are useful for producing a
CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
CC red coloured fleece. They are useful for producing coloured plant
CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
CC uses include transducing or intensifying an image, providing additional
CC light for growing phototropic organisms e.g. algae and/or corals, for
CC coating materials that experience UV damage e.g. plastics and car
CC upholstery. CFMs are useful in the flower industry, in the development of
CC new varieties of flowering plants. Other contemplated uses include,
CC expression markers, general reporter molecules, photon traps, UV sinks or
CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
CC fungal species, and in fruits and vegetables to enhance their
CC marketability. CFMs embedded in a gel matrix improve image quality in

CC situations of distorted light spectra (biomatrix). The first all-protein
CC chromophore to be isolated was Green Fluorescent protein (GFP). The
CC sequences given in records ABP69924-ABP70048 represent CFM related amino
CC acid sequences
XX SQ Sequence 223 AA;
Query Match 91.2%; Score 1156; DB 5; Length 223;
Best Local Similarity 96.4%; Pred. No. 1.7e-116;
Matches 212; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 1 SVIAKQMTYKYVMSGTVNGHYFEVEGDKGKPGYEGQTVRLAVTKGGPLPFAWDILSPQC 60
Db |||||
QY 2 SVIAKQMTYKYVMSGTVNGHYFEVEGDKGKPGYEGQTVRLAVTKGGPLPFAWDILSPQS 61
Db |||||
QY 61 QYGSIPFTKYPEDIPDYVQKSFPGRYTWERIMNFEDGAVCTVSNDSISQGNCFIYHVKFS 120
Db |||||
QY 62 QYGSIPFTKYPEDIPDYVQKSFPGRYTWERIMNFEDGAVCTVSNDSISQGNCFIYHVKFS 121
Db |||||
QY 121 GLNFPNGPVNMQKTQGWEPNTERLFARDGMLIGNNFMALKEGGHYLCEFKSTYKARK 180
Db |||||
QY 181 PVKMPGHHYVDRKLDVTNNHNDYTSVEQREISIAKPLVA 220
Db |||||
QY 182 PVKMPGHHYVDRKLDVTNNHNDYTSVEQREISIAKPPVA 221

Search completed: August 12, 2004, 06:17:05
Job time : 88.5717 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 06:19:43 ; Search time 486.141 Seconds
(without alignments)
149.169 Million cell updates/sec

Title: US-09-890-463-3
Perfect score: 1268
Sequence: 1 SVIAKQMTYKVMSTGVNGH.....STARKPLVACCFRVKSRHK 231

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1292805 seqs, 313927144 residues

Total number of hits satisfying chosen parameters: 1292805

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

1:	/cgn2_6/prodata/2/pubpa/US07_PUBCOMB.pep.*
2:	/cgn2_6/prodata/2/pubpa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/prodata/2/pubpa/US06_NEW_PUB.pep.*
4:	/cgn2_6/prodata/2/pubpa/US06_PUBCOMB.pep.*
5:	/cgn2_6/prodata/2/pubpa/US07_NEW_PUB.pep.*
6:	/cgn2_6/prodata/2/pubpa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/prodata/2/pubpa/US08_NEW_PUB.pep.*
8:	/cgn2_6/prodata/2/pubpa/US08_PUBCOMB.pep.*
9:	/cgn2_6/prodata/2/pubpa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/prodata/2/pubpa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/prodata/2/pubpa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/prodata/2/pubpa/US09_NEW_PUB.pep.*
13:	/cgn2_6/prodata/2/pubpa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/prodata/2/pubpa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/prodata/2/pubpa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/prodata/2/pubpa/US10_NEW_PUB.pep.*
17:	/cgn2_6/prodata/2/pubpa/US60_NEW_PUB.pep.*
18:	/cgn2_6/prodata/2/pubpa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	785	61.9	225	15	US-10-442-148A-7
2	785	61.9	225	15	US-10-442-148A-8
3	782	61.7	239	15	US-10-315-920-6
4	781	61.6	225	14	US-10-121-258-20
5	781	61.6	225	14	US-10-315-920-4
6	781	61.6	236	16	US-10-314-936-2
7	781	61.6	236	16	US-10-314-936-4
8	780	61.5	225	9	US-09-999-745-67
9	780	61.5	225	10	US-09-866-538-12
10	780	61.5	225	10	US-09-794-308-12
11	780	61.5	225	10	US-09-865-291-12
12	780	61.5	225	12	US-10-132-067-4
13	780	61.5	225	13	US-10-006-922-12
14	780	61.5	225	14	US-10-081-864-8
15	780	61.5	225	14	US-10-121-258-1

16	780	61.5	225	14	US-10-315-920-2
17	780	61.5	225	15	US-10-370-570-56
18	780	61.5	225	15	US-10-406-618-32
19	780	61.5	225	16	US-10-433-640-13
20	780	61.5	240	14	US-10-152-296-2
21	780	61.5	240	16	US-10-739-656-2
22	780	61.5	487	15	US-10-343-977-1
23	780	61.5	506	15	US-10-343-977-2
24	780	61.5	545	14	US-10-214-932-52
25	780	61.5	547	15	US-10-343-977-3
26	780	61.5	548	14	US-10-214-932-76
27	775	61.1	225	16	US-10-423-688A-41
28	773	61.0	227	13	US-10-006-922-10
29	767	60.5	226	14	US-10-121-258-6
30	766	60.4	225	14	US-10-121-258-24
31	765	60.3	225	13	US-10-006-922-44
32	765	60.3	225	14	US-10-081-864-12
33	765	60.3	225	14	US-10-121-258-4
34	761.5	60.1	226	16	US-10-724-178-12
35	750	59.1	230	13	US-10-006-922-18
36	750	59.1	230	14	US-10-161-403-40
37	737.5	58.2	232	14	US-10-133-973-5
38	737.5	58.2	232	15	US-10-370-570-64
39	729.5	57.5	232	16	US-10-724-178-11
40	714	56.3	225	14	US-10-121-258-8
41	714	56.3	225	16	US-10-724-178-16
42	703	55.4	225	15	US-10-370-570-61
43	687	54.2	205	13	US-10-006-922-46
44	670	52.8	231	14	US-10-133-973-6
45	670	52.8	266	13	US-10-006-922-4

ALIGNMENTS

RESULT 1
US-10-442-148A-7
; Sequence 7, Application US/10442148A
; Publication No. US20040014242A1
; GENERAL INFORMATION:
; APPLICANT: IWAKURA, MASAHIRO
; APPLICANT: HIROTA, KIYONORI
; TITLE OF INVENTION: PROCESS FOR IMMOBILIZING ORIENTATION-CONTROLLED PROTEIN AND
; METHOD OF INVENTION: PROCESS FOR ARRAYING AND IMMOBILIZING PROTEIN USING THE SAME
; FILE REFERENCE: 04583 0103-00000
; CURRENT APPLICATION NUMBER: US/10/442,148A
; PRIOR FILING DATE: 2003-05-21
; PRIOR APPLICATION NUMBER: JP 2002-148950
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic protein sequence
US-10-442-148A-7

Query Match	61.9%;	Score	785;	DB	15;	Length	225;
Best Local Similarity	62.6%;	Pred. No.	1.3e-76;				
Matches	144;	Conservative	28;	Mismatches	44;	Indels	14;
Gaps	1;						
Qy	1	SVIAKQMTYKVMSTGVNGHYEVEGDKGKPYEGEOTVRLAVTKGGLPFAWDILSPQC	60				
Db	6	NVITEFRFKVRMEGTVNGHFEFEIGEGRPVTKLVKTKGGLPFAWDILSPQF	65				
Qy	61	OYGSIPFKYPEDIPDVVKQSPGCRYTWERIMNFEDGAVCTVSNDSISQGNCFYIHKFS	120				
Db	66	OYGSKVYVKKPADIPDYKKLSFPFGFKWRVNFEDGGVAIVTQDSSIQDCFIYKVKFI	125				
Qy	121	GLNFPNGPVNMQKTOGWEPNTERLFARDGMLIGNNFMALKEGGHYLCFEKFTYKARK	180				

```
Db 126 GVNPPSDGPMVKKTGWEASTERLYPRDGLVKGETHKALKKLDGGHYLVEFKSIYMAKK 185
QY 181 PVKMPGYHYVDRKLDVTNHNKDYTSVEQREISIAKPLVACCFPRVKSRH 230
Db 186 PVQLPGYIYVDKLDITSHNEDYIVEQYE-----RTEGRH 221

RESULT 2
US-10-442-148A-8
; Sequence 8, Application US/10442148A
; Publication No. US20040014242A1
; GENERAL INFORMATION:
; APPLICANT: IWAKURA, MASAHIRO
; APPLICANT: HIROTA, KIYONORI
; TITLE OF INVENTION: PROCESS FOR IMMOBILIZING ORIENTATION-CONTROLLED PROTEIN AND
; FILE REFERENCE: 04583, 0103-00000
; CURRENT APPLICATION NUMBER: US/10/442,148A
; CURRENT FILING DATE: 2003-05-21
; PRIOR APPLICATION NUMBER: JP 2002-148950
; PRIOR FILING DATE: 2002-05-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic protein sequence
US-10-442-148A-8

Query Match 61.9%; Score 785; DB 15; Length 239;
Best Local Similarity 62.6%; Pred. No. 1.4e-76;
Matches 144; Conservative 28; Mismatches 44; Indels 14; Gaps 1;

QY 1 SVIAKQMTYKVYMSGTVNGHYFEVGDGKPKYEGEQTVRLAVTKGGLPFAWDILSPQC 60
Db 6 NVITEPMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVKTKGGLPFAWDILSPQF 65
QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFEDGAVCTVSDSSIQNCFCFIYHKFS 120
Db 66 QYGSKVYVHKPADIPDYKLSFEGFKWERNVNFEDGGVATVTDSSLQDGCFCFIYKVKFI 125
QY 121 GLNFPPNGPMVKKTQGWEPNTERLFARDGMLIGNNFMAKLEGGHYLCEFKSTYKARK 180
Db 126 GVNFPDGPVMQKKTGWEASTERLYPRDGLVKGETHKALKKLDGGHYLVEFKSIYMAKK 185
QY 181 PVKMPGYHYVDRKLDVTNHNKDYTSVEQREISIAKPLVACCFPRVKSRH 230
Db 186 PVQLPGYIYVDKLDITSHNEDYIVEQYE-----RTEGRH 221

RESULT 3
US-10-315-920-6
; Sequence 6, Application US/10315920
; Publication No. US20030175809A1
; GENERAL INFORMATION:
; APPLICANT: Fradkov, Arcady Fedorovich
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
; FILE REFERENCE: CLON-077CIP
; CURRENT APPLICATION NUMBER: US/10/315,920
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/211,607
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: PCT/US01/19097
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic protein sequence
US-10-315-920-6

Query Match 61.9%; Score 785; DB 14; Length 225;
Best Local Similarity 62.6%; Pred. No. 3.5e-76;
Matches 144; Conservative 28; Mismatches 44; Indels 14; Gaps 1;

QY 1 SVIAKQMTYKVYMSGTVNGHYFEVGDGKPKYEGEQTVRLAVTKGGLPFAWDILSPQC 60
Db 6 NVITEPMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVKTKGGLPFAWDILSPQF 65
QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFEDGAVCTVSDSSIQNCFCFIYHKFS 120
Db 66 QYGSKVYVHKPADIPDYKLSFEGFKWERNVNFEDGGVATVTDSSLQDGCFCFIYKVKFI 125
QY 121 GLNFPPNGPMVKKTQGWEPNTERLFARDGMLIGNNFMAKLEGGHYLCEFKSTYKARK 180
Db 126 GVNFPDGPVMQKKTGWEASTERLYPRDGLVKGETHKALKKLDGGHYLVEFKSIYMAKK 185
QY 181 PVKMPGYHYVDRKLDVTNHNKDYTSVEQREISIAKPLVACCFPRVKSRH 230
Db 186 PVQLPGYIYVDKLDITSHNEDYIVEQYE-----RTEGRH 221
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: variant of sequence from Discosoma sp.
US-10-315-920-6

Query Match 61.7%; Score 782; DB 14; Length 225;
Best Local Similarity 62.6%; Pred. No. 2.7e-76;
Matches 144; Conservative 28; Mismatches 44; Indels 14; Gaps 1;

QY 1 SVIAKQMTYKVYMSGTVNGHYFEVGDGKPKYEGEQTVRLAVTKGGLPFAWDILSPQC 60
Db 6 NVITEPMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVKTKGGLPFAWDILSPQF 65
QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFEDGAVCTVSDSSIQNCFCFIYHKFS 120
Db 66 QYGSKVYVHKPADIPDYKLSFEGFKWERNVNFEDGGVATVTDSSLQDGCFCFIYKVKFI 125
QY 121 GLNFPPNGPMVKKTQGWEPNTERLFARDGMLIGNNFMAKLEGGHYLCEFKSTYKARK 180
Db 126 GVNFPDGPVMQKKTGWEASTERLYPRDGLVKGETHKALKKLDGGHYLVEFKSIYMAKK 185
QY 181 PVKMPGYHYVDRKLDVTNHNKDYTSVEQREISIAKPLVACCFPRVKSRH 230
Db 186 PVQLPGYIYVDKLDITSHNEDYIVEQYE-----RTEGRH 221

RESULT 4
US-10-121-258-20
; Sequence 20, Application US/10121258
; Publication No. US20030059835A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; FILE REFERENCE: UC083.1CP2CP1
; CURRENT APPLICATION NUMBER: US/10/121,258
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DsRed with I125R
US-10-121-258-20

Query Match 61.6%; Score 781; DB 14; Length 225;
Best Local Similarity 62.6%; Pred. No. 3.5e-76;
Matches 144; Conservative 28; Mismatches 44; Indels 14; Gaps 1;

QY 1 SVIAKQMTYKVYMSGTVNGHYFEVGDGKPKYEGEQTVRLAVTKGGLPFAWDILSPQC 60
Db 6 NVITEPMRFKVRMEGTVNGHEFEIEGEGEGRPYEGHNTVKLVKTKGGLPFAWDILSPQF 65
QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFEDGAVCTVSDSSIQNCFCFIYHKFS 120
Db 66 QYGSKVYVHKPADIPDYKLSFEGFKWERNVNFEDGGVATVTDSSLQDGCFCFIYKVKFI 125
QY 121 GLNFPPNGPMVKKTQGWEPNTERLFARDGMLIGNNFMAKLEGGHYLCEFKSTYKARK 180
Db 126 GVNFPDGPVMQKKTGWEASTERLYPRDGLVKGETHKALKKLDGGHYLVEFKSIYMAKK 185
QY 181 PVKMPGYHYVDRKLDVTNHNKDYTSVEQREISIAKPLVACCFPRVKSRH 230
Db 186 PVQLPGYIYVDKLDITSHNEDYIVEQYE-----RTEGRH 221
```

```
RESULT 5
US-10-315-920-4
; Sequence 4, Application US/10315920
; Publication No. US20030175809A1
; GENERAL INFORMATION:
; APPLICANT: Fradkov, Arcady Fedorovich
; APPLICANT: Tersikh, Alexey
; TITLE OF INVENTION: FLUORESCENT TIMER PROTEINS AND METHODS
; TITLE OF INVENTION: FOR THEIR USE
; FILE REFERENCE: CLON-077CIP
; CURRENT APPLICATION NUMBER: US/10/315,920
; PRIOR FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: 60/211,607
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: PCT/US01/19097
; PRIOR FILING DATE: 2001-06-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: variant of sequence from Discosoma sp.
US-10-315-920-4

Query Match          61.6%; Score 781; DB 14; Length 225;
Best Local Similarity 62.6%; Pred. No. 3.5e-76;
Matches 144; Conservative 28; Mismatches 44; Indels 14; Gaps 1;

QY 1 SVIAKQMTYKVMGSGTVNGHYFEVGDGKPKYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 6 NVIKFMRFKVMRMEGTVMNGHEFEIEGEGEGRPYEGHNTVKLVKTKGGPLPFAWDILSPQF 65

QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFEDGAVCTVSNDSISQNCFIYHVKFS 120
   ||||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 66 QYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVATVTQDSSLQDGCIFYVKVFI 125

QY 121 GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGHYLCPEKSTYKARK 180
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 126 GYNFSDGPMQKTMGWAEASTERLYPRDGLVKGELHKAALKDKDGGHYLVEFKTIYMAKK 185

QY 181 PVKMPGYHYVDRKLDVTNHNKDYTSVEQREISIAKPLVACCFRVRKSRH 230
   ||||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 186 PVQLPGYVYVDSKLDITSHNKDYTIVEQYE-----RTEGRH 221

RESULT 7
US-10-314-936-4
; Sequence 4, Application US/10314936
; Publication No. US20040110225A1
; GENERAL INFORMATION:
; APPLICANT: Gibbs, Patrick D.L.
; APPLICANT: Carter, Robert W.
; APPLICANT: Schmale, Michael C.
; TITLE OF INVENTION: FLUORESCENT PROTEINS FROM AQUATIC SPECIES
; FILE REFERENCE: 638.004
; CURRENT APPLICATION NUMBER: US/10/314,936
; CURRENT FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 236
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: mutant red fluorescent protein
US-10-314-936-4

Query Match          61.6%; Score 781; DB 16; Length 236;
Best Local Similarity 62.6%; Pred. No. 3.7e-76;
Matches 144; Conservative 28; Mismatches 44; Indels 14; Gaps 1;

QY 1 SVIAKQMTYKVMGSGTVNGHYFEVGDGKPKYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 6 NVIKFMRFKVMRMEGTVMNGHEFEIEGEGEGRPYEGHNTVKLVKTKGGPLPFAWDILSPQF 65

QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNFEDGAVCTVSNDSISQNCFIYHVKFS 120
   ||||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 66 QYGSKVYVVKHPADIPDYKKLSFPEGFKWERVMNFEDGGVATVTQDSSLQDGCIFYVKVFI 125

QY 121 GLNFPNGPVMQKKTQGWEPNTERLFARDGMLIGNNFMALKLEGGHYLCPEKSTYKARK 180
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 126 GYNFSDGPMQKTMGWAEASTERLYPRDGLVKGELHKAALKDKDGGHYLVEFKTIYMAKK 185

QY 181 PVKMPGYHYVDRKLDVTNHNKDYTSVEQREISIAKPLVACCFRVRKSRH 230
   ||||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 186 PVQLPGYVYVDSKLDITSHNKDYTIVEQYE-----RTEGRH 221

RESULT 8
US-09-999-745-67
; Sequence 67, Application US/09999745
; Patent No. US20020157120A1
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Baird, Geoffrey
; TITLE OF INVENTION: CIRCULARLY PERMUTED FLUORESCENT PROTEIN INDICATORS
; FILE REFERENCE: REGEN1470-1
; CURRENT APPLICATION NUMBER: US/09/999,745
; CURRENT FILING DATE: 2001-10-23
; PRIOR APPLICATION NUMBER: 09/316,920
```

```
; PRIOR FILING DATE: 1999-05-21
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 67
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-09-999-745-67

Query Match          61.5%; Score 780; DB 9; Length 225;
Best Local Similarity 62.6%; Pred. No. 4.4e-76;
Matches 144; Conservative 28; Mismatches 44; Indels 14; Gaps 1;

QY 1 SVIAKQMTYKVMGTVNGHYFEVEGDGKPKYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 6 NVIKFMRFKVRMEGTVNGHFEIEGEGRPYEGHNTVKLVTKGGPLPFAWDILSPQF 65

QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHKFS 120
Db 66 QYGSKVYVVKHPADIPDYKLSFPEGFKWERVMNFDGGVTVTQDSSLQDGCFFIYKVKFI 125

QY 121 GLNFPNGPVMOKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCPEKSTYKARK 180
Db 126 GVNFDSDGVPVMOKKTGWGEASTERLYPRDGLVKGELHKAALKDKDGHYLVPEKSIYMAKK 185

QY 181 PVKMPGYHYVDRKLDVTNHNKDYTSVEQREISIAARKPLVACCFRVRKSRH 230
Db 186 PVQLPGYVYVDSKLDITSHNEDYTVIVEQVE-----RTEGRH 221
```

```
RESULT 9
US-09-866-538-12
; Sequence 12, Application US/09866538
; Publication No. US20030032088A1
; GENERAL INFORMATION:
; APPLICANT: TSJEN, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
; FILE REFERENCE: REGEN1530-2
; CURRENT APPLICATION NUMBER: US/09/866,538
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-09-866-538-12

Query Match          61.5%; Score 780; DB 10; Length 225;
Best Local Similarity 62.6%; Pred. No. 4.4e-76;
Matches 144; Conservative 28; Mismatches 44; Indels 14; Gaps 1;

QY 1 SVIAKQMTYKVMGTVNGHYFEVEGDGKPKYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 6 NVIKFMRFKVRMEGTVNGHFEIEGEGRPYEGHNTVKLVTKGGPLPFAWDILSPQF 65

QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHKFS 120
Db 66 QYGSKVYVVKHPADIPDYKLSFPEGFKWERVMNFDGGVTVTQDSSLQDGCFFIYKVKFI 125

QY 121 GLNFPNGPVMOKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCPEKSTYKARK 180
Db 126 GVNFDSDGVPVMOKKTGWGEASTERLYPRDGLVKGELHKAALKDKDGHYLVPEKSIYMAKK 185

QY 181 PVKMPGYHYVDRKLDVTNHNKDYTSVEQREISIAARKPLVACCFRVRKSRH 230
Db 186 PVQLPGYVYVDSKLDITSHNEDYTVIVEQVE-----RTEGRH 221
```

```
RESULT 10
US-09-794-308-12
```

```
; Sequence 12, Application US/09794308
; Publication No. US20030170911A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSJEN, Roger
; APPLICANT: ZACHARIAS, David
; APPLICANT: BAIRD, Geoffrey
; TITLE OF INVENTION: NON-OLIGOMERIZING FLUORESCENT PROTEINS
; FILE REFERENCE: REGEN1530
; CURRENT APPLICATION NUMBER: US/09/794,308
; CURRENT FILING DATE: 2001-02-26
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-09-794-308-12
```

```
Query Match          61.5%; Score 780; DB 10; Length 225;
Best Local Similarity 62.6%; Pred. No. 4.4e-76;
Matches 144; Conservative 28; Mismatches 44; Indels 14; Gaps 1;

QY 1 SVIAKQMTYKVMGTVNGHYFEVEGDGKPKYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 6 NVIKFMRFKVRMEGTVNGHFEIEGEGRPYEGHNTVKLVTKGGPLPFAWDILSPQF 65

QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHKFS 120
Db 66 QYGSKVYVVKHPADIPDYKLSFPEGFKWERVMNFDGGVTVTQDSSLQDGCFFIYKVKFI 125

QY 121 GLNFPNGPVMOKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCPEKSTYKARK 180
Db 126 GVNFDSDGVPVMOKKTGWGEASTERLYPRDGLVKGELHKAALKDKDGHYLVPEKSIYMAKK 185

QY 181 PVKMPGYHYVDRKLDVTNHNKDYTSVEQREISIAARKPLVACCFRVRKSRH 230
Db 186 PVQLPGYVYVDSKLDITSHNEDYTVIVEQVE-----RTEGRH 221
```

```
RESULT 11
US-09-865-291-12
; Sequence 12, Application US/09865291
; Publication No. US20030186229A1
; GENERAL INFORMATION:
; APPLICANT: REGENTS OF THE UNIVERSITY OF CALIFORNIA
; APPLICANT: TSJEN, Roger
; APPLICANT: TING, Alice
; APPLICANT: ZHANG, Jin
; TITLE OF INVENTION: EMISSION RATIO-METRIC INDICATORS OF PHOSPHORYLATION
; FILE REFERENCE: REGEN1550
; CURRENT APPLICATION NUMBER: US/09/865,291
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
US-09-865-291-12
```

```
Query Match          61.5%; Score 780; DB 10; Length 225;
Best Local Similarity 62.6%; Pred. No. 4.4e-76;
Matches 144; Conservative 28; Mismatches 44; Indels 14; Gaps 1;
```

```
QY 1 SVIAKQMTYKVMGTVNGHYFEVEGDGKPKYEGEQTVRLAVTKGGPLPFAWDILSPQC 60
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 6 NVIKFMRFKVRMEGTVNGHFEIEGEGRPYEGHNTVKLVTKGGPLPFAWDILSPQF 65

QY 61 QYGSIPFTKYPEDIPDYVKQSPFGRYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHKFS 120
Db 66 QYGSKVYVVKHPADIPDYKLSFPEGFKWERVMNFDGGVTVTQDSSLQDGCFFIYKVKFI 125
```

QY	121	GLNPPNGPVWQKKTQTGBENTERLFAEDGMLIGNNFMAKLKEGGHYLCFFKSTYKARK	180
Db	126	GVNFPDGPVWQKKTGMGEASTERYLPDRGVLGKEIHKALKLKDGGHLYVEFKSIYMAKK	185
QY	181	PVKMPGYHYVDRLDVTNNHKNKYTSVEOREISIAARKPLVACCFPRVKSRH	230
Db	186	PVQLPGYYVDSKLDITSHNEDYTIVEQYE-----RTEGRH	221
 RESULT 12			
US-10-132-067-4			
; Sequence 4, Application US/10132067			
; Publication No. US20030203355A1			
; GENERAL INFORMATION:			
; APPLICANT: Bradbury, Andrew			
; APPLICANT: Zeytun, Ahmet			
; APPLICANT: Waldo, Geoffrey			
; APPLICANT: The Regents of the University of California			
; TITLE OF INVENTION: Fluorobodies: Binding Ligands With Intrinsic			
; FILE REFERENCE: 021362-000600US			
; CURRENT FILING DATE: 2002-04-24			
; NUMBER OF SEQ ID NOS: 48			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 4			
; LENGTH: 225			
; TYPE: PRT			
; ORGANISM: Discosoma sp.			
; FEATURES:			
; OTHER INFORMATION: red fluorescent protein (dsRED)			
US-10-132-067-4			
 Query Match 61.5%; Score 780; DB 12; Length 225;			
Best Local Similarity 62.6%; Pred. No. 4.4e-76;			
Matches 144; Conservative 28; Mismatches 44; Indels 14; Gaps 19			
QY	1	SVIAKQMTYKVYMSGTVNGHYFEVGDKGPKPYEGEQTVRLAVTKGGPLPFAWDILSPQC	60
Db	6	NVIKEPFKVRMEGTVNGHEFEIEGEGRPYECHNTVKLVKTKGGPLPFAWDILSPQF	65
QY	61	QYGSIPTKYPEDIPDYVKQSPGPRYWTERIMNPEDGAVCTVSNDSSIQGNCFIYHVXFS	120
Db	66	QYGSKVYVWKHPADIPDYKLSFPEGFKWERVMNFDGGVTVTQDSLQDGCFIYKVKFI	125
QY	121	GLNPPNGPVWQKKTQTGBENTERLFAEDGMLIGNNFMAKLKEGGHYLCFFKSTYKARK	180
Db	126	GVNFPDGPVWQKKTGMGEASTERYLPDRGVLGKEIHKALKLKDGGHLYVEFKSIYMAKK	185
QY	181	PVKMPGYHYVDRLDVTNNHKNKYTSVEOREISIAARKPLVACCFPRVKSRH	230
Db	186	PVQLPGYYVDSKLDITSHNEDYTIVEQYE-----RTEGRH	221
 RESULT 13			
US-10-006-922-12			
; Sequence 12, Application US/10006922			
; Publication No. US20020197676A1			
; GENERAL INFORMATION:			
; APPLICANT: Lukanov, Sergey A			
; APPLICANT: Fradkov, Arcady F.			
; APPLICANT: Labas, Yulii A.			
; APPLICANT: Matz, Mikhail V.			
; APPLICANT: Tersikh, Alexey			
; TITLE OF INVENTION: No. US20020197676A1el Chromophores/Fluorophores and			
; FILE REFERENCE: CLON-035CIP			
; CURRENT APPLICATION NUMBER: US/10/006,922			
; CURRENT FILING DATE: 2001-12-04			
; PRIOR APPLICATION NUMBER: 09/120,330			
; PRIOR FILING DATE: 1998-12-11			
; PRIOR APPLICATION NUMBER: 09/457,898			
; PRIOR FILING DATE: 1999-12-09			

Db 66 QYGSKVYVKKHPADIPDYKLSFPEGKWERVMNFEDGGVTVTQDSSLQDGCIFYKVKFI 125
QY 121 GLNFPNGPVMOKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCBFKSTYKARK 180
Db 126 GVNFPDGPVMOKKTGWGEASTERLYPRDGVLYKGEIHKALKDKGGHYLVEFKSIYMAKK 185
QY 181 PVKMPGYHYVDRKLDVTNNKDYTSVEQREISIAKPKLVACCFRVRKSRH 230
Db 186 PVQLPGYVYVDSKLDITSHNEDYTIIVEQYE-----RTEGRH 221

RESULT 15
US-10-121-258-1
; Sequence 1, Application US/10121258
; Publication No. US20030059835A1
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger
; APPLICANT: Campbell, Robert
; TITLE OF INVENTION: MONOMERIC AND DIMERIC FLUORESCENT
; FILE REFERENCE: UC083.1CP2CP1
; CURRENT APPLICATION NUMBER: US/10/121,258
; CURRENT FILING DATE: 2002-04-10
; PRIOR APPLICATION NUMBER: 09/794,308
; PRIOR FILING DATE: 2001-02-26
; PRIOR APPLICATION NUMBER: 09/866,538
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Discosoma sp.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(225)
; OTHER INFORMATION: wild-type DeRed
US-10-121-258-1

Query Match 61.5%; Score 780; DB 14; Length 225;
Best Local Similarity 62.6%; Pred No. 4.4e-76;
Matches 144; Conservative 28; Mismatches 44; Indels 14; Gaps 1;
QY 1 SVIAKQMTYKVMSTVNGHYFVEVGDGKPYEGEQTVRVLAVTKGGPLPFAWDILSPQC 60
Db 6 NVIKBFMEFKVRMEGTVNGHEFEIEGEGERPYEGHNTVKLVTKGGPLPFAWDILSPQF 65
QY 61 QYGSIPFTKYPEDIPDYVYKQSPGPGRYTWERIMNPFEDGAVCTVSDSSIQGNCFIYHVKFS 120
Db 66 QYGSKVYVKKHPADIPDYKLSFPEGKWERVMNFEDGGVTVTQDSSLQDGCIFYKVKFI 125
QY 121 GLNFPNGPVMOKKTQGWEPNTERLFARDGMLIGNFMALKLEGGHYLCBFKSTYKARK 180
Db 126 GVNFPDGPVMOKKTGWGEASTERLYPRDGVLYKGEIHKALKDKGGHYLVEFKSIYMAKK 185
QY 181 PVKMPGYHYVDRKLDVTNNKDYTSVEQREISIAKPKLVACCFRVRKSRH 230
Db 186 PVQLPGYVYVDSKLDITSHNEDYTIIVEQYE-----RTEGRH 221

Search completed: August 12, 2004, 06:51:21
Job time : 487.141 secs

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	43	51.8	107	4	US-09-106-568E-43	Sequence 43, Appl
2	39	47.0	107	4	US-09-106-568E-80	Sequence 80, Appl
3	39	47.0	3665	2	US-08-222-617A-13	Sequence 13, Appl
4	39	47.0	3712	2	US-08-222-617A-4	Sequence 4, Appl
5	39	47.0	3712	2	US-08-222-617A-25	Sequence 25, Appl
6	38	45.8	195	4	US-09-489-039A-13455	Sequence 13455, A
7	38	45.8	209	4	US-09-134-001C-4478	Sequence 4478, A
8	37.5	45.2	329	4	US-09-540-236-3211	Sequence 3211, Ap
9	37	44.6	171	4	US-09-540-236-2461	Sequence 2461, Ap
10	37	44.6	351	4	US-09-198-452A-991	Sequence 991, App
11	36	43.4	390	4	US-09-489-039A-13547	Sequence 13547, A
12	36	43.4	431	4	US-09-107-532A-7056	Sequence 7056, Ap
13	35	42.2	28	1	US-08-446-692-3	Sequence 3, Appl
14	35	42.2	28	2	US-08-488-351A-3	Sequence 3, Appl
15	35	42.2	28	4	US-09-106-568E-21	Sequence 21, Appl
16	35	42.2	30	3	US-09-100-409A-42	Sequence 42, Appl
17	35	42.2	30	5	PCT-US95-13841-6	Sequence 6, Appl
18	35	42.2	38	1	US-08-446-692-11	Sequence 11, Appl
19	35	42.2	38	2	US-08-488-351A-11	Sequence 11, Appl
20	35	42.2	87	4	US-09-134-000C-5981	Sequence 5981, Ap
21	35	42.2	104	2	US-08-292-968-23	Sequence 23, Appl
22	35	42.2	104	2	US-08-467-974-23	Sequence 23, Appl
23	35	42.2	104	2	US-08-467-536-23	Sequence 23, Appl
24	35	42.2	104	3	US-08-467-976-23	Sequence 23, Appl
25	35	42.2	104	3	US-09-082-514-23	Sequence 23, Appl
26	35	42.2	107	4	US-09-106-568E-60	Sequence 60, Appl
27	35	42.2	174	4	US-09-252-991A-25897	Sequence 25897, A

SOFTWARE: Microsoft Word 97
SEQ ID NO 80
LENGTH: 107
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Polypeptide segment of ACVS_CEPAC shown in Figure 4.
US-09-106-569E-80

Query Match 47.0%; Score 39; DB 4; Length 107;
Best Local Similarity 53.8%; Pred. No. 5.7;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 AKQMTYKVMYSGT 16
:|:|:|:|:|:|
Db 4 SKQLAYVTYTSQT 16

RESULT 3

US-08-222-617A-13
Sequence 13, Application US/08222617A
Patent No. 5882879

GENERAL INFORMATION:

APPLICANT: Veenstra, Annemarie E.
APPLICANT: Martin, Juan F.
APPLICANT: Garcia, Bruno D.
APPLICANT: Gutierrez, Santiago
APPLICANT: Barredo, Jose L.
APPLICANT: Von Doehren, Hans
APPLICANT: Palissa, Harriet
APPLICANT: Van Liempt, Henk
APPLICANT: Montenegro, Eduardo P.
TITLE OF INVENTION: A Method for Influencing Beta-Lactam
TITLE OF INVENTION: Antibiotic Production and for Isolation of Large
TITLE OF INVENTION: Quantities of ACV Synthetase
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/222,617A
FILING DATE: 04-APR-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

REFERENCE/DOCKET NUMBER: 97,157

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 3665 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

ORGANISM: Acetomoniun chrysogenum

FEATURE:

NAME/KEY: Protein

LOCATION: 1..3665

OTHER INFORMATION: /label= ACVS

OTHER INFORMATION: /note= "ACV Synthetase from Acetomoniun

OTHER INFORMATION: chrysogenum; aa 1-3665"

US-08-222-617A-13

Query Match 47.0%; Score 39; DB 2; Length 3665;
Best Local Similarity 53.8%; Pred. No. 3.4e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 AKQMTYKVMYSGT 16
:|:|:|:|:|:|
Db 414 SKQLAYVTYTSQT 426

RESULT 4

US-08-222-617A-4
Sequence 4, Application US/08222617A
Patent No. 5882879

GENERAL INFORMATION:

APPLICANT: Veenstra, Annemarie E.
APPLICANT: Martin, Juan F.
APPLICANT: Garcia, Bruno D.
APPLICANT: Gutierrez, Santiago
APPLICANT: Barredo, Jose L.
APPLICANT: Von Doehren, Hans
APPLICANT: Palissa, Harriet
APPLICANT: Van Liempt, Henk
APPLICANT: Montenegro, Eduardo P.
TITLE OF INVENTION: A Method for Influencing Beta-Lactam
TITLE OF INVENTION: Antibiotic Production and for Isolation of Large
TITLE OF INVENTION: Quantities of ACV Synthetase
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/222,617A

FILING DATE: 04-APR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

REFERENCE/DOCKET NUMBER: 97,157

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 3712 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

FEATURE:

NAME/KEY: Modified-site

LOCATION: 2555

OTHER INFORMATION:

OTHER INFORMATION: /note= "Xaa=Ala or Ser"

US-08-222-617A-4

Query Match 47.0%; Score 39; DB 2; Length 3712;
Best Local Similarity 53.8%; Pred. No. 3.4e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 AKQMTYKVMYSGT 16
:|:|:|:|:|:|
Db 414 SKQLAYVTYTSQT 426

RESULT 5

US-08-222-617A-25
Sequence 25, Application US/08222617A
Patent No. 5882879

GENERAL INFORMATION:

APPLICANT: Veenstra, Annemarie E.

APPLICANT: Martin, Juan F.
APPLICANT: Garcia, Bruno D.
APPLICANT: Gutierrez, Santiago
APPLICANT: Barredo, Jose L.
APPLICANT: Von Doehren, Hans
APPLICANT: Palissa, Harriet
APPLICANT: Van Liempt, Henk
APPLICANT: Montenegro, Eduardo P.
TITLE OF INVENTION: A Method for Influencing Beta-Lactam
TITLE OF INVENTION: Antibiotic Production and for Isolation of Large
TITLE OF INVENTION: Quantities of ACV Synthetase
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/222,617A
FILING DATE: 04-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
REFERENCE/DOCKET NUMBER: 97,157
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 3712 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-222-617A-25

Query Match 47.0%; Score 39; DB 2; Length 3712;
Best Local Similarity 53.8%; Pred. No. 3.4e+02;
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 AKQMTYKVMSTG 16
Db 414 SKQLAYVYTSQT 426

RESULT 6
US-09-489-039A-13455
Sequence 13455, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 13455
LENGTH: 195
TYPE: PRT
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13455

Query Match 45.8%; Score 38; DB 4; Length 195;
Best Local Similarity 57.1%; Pred. No. 18;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 VIAQMTYKVMMSG 15
| : | | | : | |

Db 86 VLEKQMTSRGYMIG 99

RESULT 7
US-09-134-001C-4478
Sequence 4478, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4478
LENGTH: 209
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4478

Query Match 45.8%; Score 38; DB 4; Length 209;
Best Local Similarity 70.0%; Pred. No. 19;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 8 TYKVYMSGTV 17
| | | | : | |
Db 93 TYKVYKNGV 102

RESULT 8
US-09-540-236-3211
Sequence 3211, Application US/09540236
Patent No. 6673910
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARI
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2005-001
CURRENT APPLICATION NUMBER: US/09/540,236
CURRENT FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 3840
SEQ ID NO 3211
LENGTH: 329
TYPE: PRT
ORGANISM: M.catarrhalis
US-09-540-236-3211

Query Match 45.2%; Score 37.5; DB 4; Length 329;
Best Local Similarity 41.2%; Pred. No. 40;
Matches 7; Conservative 8; Mismatches 1; Indels 1; Gaps 1;

QY 1 SVIAKQMTYKVMSTG 17
| : : : | | | : : : | |
Db 154 TIVAKQITTYQM-VNGNV 169

RESULT 9
US-09-540-236-2461
Sequence 2461, Application US/09540236
Patent No. 6673910
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARI
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.2005-001
CURRENT APPLICATION NUMBER: US/09/540,236
CURRENT FILING DATE: 2000-04-04
NUMBER OF SEQ ID NOS: 3840

```
; SEQ ID NO 2461
; LENGTH: 171
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2461

Query Match      44.6%; Score 37; DB 4; Length 171;
Best Local Similarity 63.6%; Pred. No. 23;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      6 QNTYKVMVSGT 16
Db      28 QKTYKVIYINH 38

RESULT 10
US-09-198-452A-991
; Sequence 991, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments thereof and uses thereof, in particular for the diagnosis, prevention and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 991
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-991

Query Match      44.6%; Score 37; DB 4; Length 351;
Best Local Similarity 50.0%; Pred. No. 53;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      3 IAKQMTYKVMVSGT 16
Db      64 LARGMTYKAIISNT 77

RESULT 11
US-09-489-039A-13547
; Sequence 13547, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13547
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13547

Query Match      43.4%; Score 36; DB 4; Length 390;
Best Local Similarity 46.7%; Pred. No. 93;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY      3 IAKQMTYKVMVSGT 17
Db      178 VAEFASQKVVYSGV 192

RESULT 12
US-09-107-532A-7056
; Sequence 7056, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 7056:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...431
; SEQUENCE DESCRIPTION: SEQ ID NO: 7056:
US-09-107-532A-7056

Query Match      43.4%; Score 36; DB 4; Length 431;
Best Local Similarity 46.7%; Pred. No. 1e+02;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY      3 IAKQMTYKVMVSGT 17
Db      292 IIKVMTYSIMLGWV 306

RESULT 13
US-08-446-692-3
; Sequence 3, Application US/08446692
; Patent No. 5759551
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
```

```

; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-446-692-3

```

```

Query Match 42.2%; Score 35; DB 1; Length 28;
Best Local Similarity 45.5%; Pred. No. 6.8;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 5 KOMTYKVYMSG 15
::: | :|||
Db 4 RLLYMYMSG 14

```

```

RESULT 14
US-08-488-351A-3
; Sequence 3, Application US/08488351A
; Patent No. 5843446
; GENERAL INFORMATION:
; APPLICANT: Ladd, Anna
; APPLICANT: Wang, Chang Yi
; APPLICANT: Zamb, Timothy
; TITLE OF INVENTION: Immunogenic LHRH peptide constructs
; TITLE OF INVENTION: and synthetic universal immune stimulators for vaccines
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Maria C.H. Lin
; STREET: 345 Park Avenue
; CITY: New York
; STATE: NY
; COUNTRY: US
; ZIP: 10154-0053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,351A
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; APPLICATION DATA:
; APPLICATION NUMBER: US 08/446,692
; FILING DATE: 7-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/229,275
; FILING DATE: 14-APR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US 08/057,166
; FILING DATE: 27-APR-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Maria C.H. Lin
; REGISTRATION NUMBER: 29,323
; REFERENCE/DOCKET NUMBER: 1151-4146 US2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)415-8745
; TELEFAX: (516)751-6849
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 28 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-488-351A-3

```

```

Query Match 42.2%; Score 35; DB 2; Length 28;
Best Local Similarity 45.5%; Pred. No. 6.8;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 5 KOMTYKVYMSG 15
::: | :|||
Db 4 RLLYMYMSG 14

```

```

RESULT 15
US-09-106-568E-21
; Sequence 21, Application US/09106568E
; Patent No. 6455248
; GENERAL INFORMATION:
; APPLICANT: Bhattacharjee, J.
; APPLICANT: Suvarna, Kalavati
; APPLICANT: Bhattacherjee, Vaskar
; TITLE OF INVENTION: METHODS AND REAGENTS FOR DETECTING FUNGAL PATHOGENS IN
; TITLE OF INVENTION: A BIOLOGICAL SAMPLE
; FILE REFERENCE: 96,247-A
; CURRENT APPLICATION NUMBER: US/09/106,568E
; CURRENT FILING DATE: 1998-06-29
; PRIOR APPLICATION NUMBER: 08/650,809
; PRIOR FILING DATE: 1997-05-20
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: Microsoft Word 97
; SEQ ID NO 21
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Polypeptide segment of ACVT_PENCH shown in Figure 4.
; US-09-106-568E-21

```

```

Query Match 42.2%; Score 35; DB 4; Length 28;
Best Local Similarity 46.2%; Pred. No. 6.8;
Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 4 AKOMTYKVYMSGT 16
::: | :|||
Db 4 SQQLAYVTYTSGT 16

```

```

Search completed: August 12, 2004, 06:21:05
Job time : 2.77664 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 07:01:35 ; Search time 37 Seconds
(without alignments)
44.196 Million cell updates/sec

Title: US-09-890-463-2

Perfect score: 83

Sequence: 1 SVIAKQMTYKVMGTV 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 2990

Minimum DB seq length: 0

Maximum DB seq length: 17

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	31.3	16	2 B60278	24K antigen - Myco
2	22	26.5	11	2 C49037	TcR gamma V-J regi
3	22	26.5	14	2 A47421	Leukotriene B-4 12
4	22	26.5	15	2 PH1616	Ig H chain V-D-J r
5	22	26.5	16	2 B45895	T-cell surface gly
6	22	26.5	17	2 S78756	ribosomal protein
7	21	25.3	12	2 C49215	urease (EC 3.5.1.5
8	21	25.3	16	2 S09732	photosystem I prot
9	21	25.3	17	2 B42965	talin (glycosylate
10	21	25.3	17	2 S69164	ferredoxin al - Ja
11	20	24.1	8	2 S63493	dissimilatory sulf
12	20	24.1	9	2 G41946	T-cell receptor ga
13	20	24.1	9	2 G85802	hypothetical prote
14	20	24.1	10	2 S66458	ferredoxin - Rhizo
15	20	24.1	12	2 H41946	T-cell receptor ga
16	20	24.1	13	2 PA0049	protein QA100046 -
17	20	24.1	14	2 PH1617	Ig H chain V-D-J r
18	20	24.1	15	2 S03353	Plastocyanin - Mic
19	20	24.1	15	2 PS0251	15K protein 5106 -
20	19	22.9	11	2 B49037	TcR gamma V-J regi
21	19	22.9	12	2 PN0162	malate dehydrogena
22	19	22.9	12	2 A37933	Ig lambda chain J
23	19	22.9	13	2 PN0125	serine proteinase
24	19	22.9	13	2 PS0325	tetrahydroberberin
25	19	22.9	14	2 A60158	disaggregatase - M
26	19	22.9	14	2 PA0101	protein QF200020 -
27	19	22.9	15	2 S21241	oligo-1,6-glucosid
28	19	22.9	15	2 S21240	alpha-glucosidase
29	19	22.9	15	2 S21202	glucan 1,4-alpha-g

30	19	22.9	15	2 PA0099	phenotypic variati
31	19	22.9	15	2 A27504	histone H2A - mous
32	19	22.9	16	2 F41299	T-cell receptor al
33	19	22.9	16	2 S24667	protein-tyrosine k
34	19	22.9	16	2 S55307	glutathione transf
35	19	22.9	17	2 A27486	folitropin inhibi
36	18	21.7	9	2 I46016	cytokeatin 4 - bo
37	18	21.7	11	2 PN0167	ribosomal protein
38	18	21.7	11	2 PT0214	T-cell receptor be
39	18	21.7	11	2 I60434	68kDa neurofilam
40	18	21.7	12	2 PH1611	Ig H chain V-D-J r
41	18	21.7	13	2 JQ2309	hypothetical 1.6K
42	18	21.7	13	2 JQ2319	hypothetical 1.6K
43	18	21.7	13	2 B20907	Ig kappa-1 chain J
44	18	21.7	15	2 A41338	isocitrate lyase (
45	18	21.7	15	2 S43321	RNA-binding protei

ALIGNMENTS

RESULT 1

B60278
24K antigen - Mycobacterium bovis (fragment)
C:Species: Mycobacterium bovis

C:Date: 11-Dec-1992 #sequence_revision 11-Dec-1992 #text_change 18-Jun-1993
C:Accession: B60278

R:Rifis, T.; Costopoulos, C.; Radford, A.J.; Bacic, A.; Wood, P.R.

Infect. Immun. 59, 800-807, 1991

A:Title: Purification and characterization of major antigens from a Mycobacterium bovis (

A:Reference number: A60278; MUID:91147217; PMID:1900061

A:Accession: B60278

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-16 <PIF>

Query Match 31.3%; Score 26; DB 2; Length 16;

Best Local Similarity 55.6%; Pred. No. 3.4e+02;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 TYKVMGSGT 16

||| : ||

DB 4 TYKEELKGT 12

RESULT 2

C49037
TcR gamma V-J region - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999

C:Accession: C49037

R:Esqueria, A.; Wilde, D.B.; McConnell, T.J.; Sturmhofel, K.; Valas, R.B.; Shevach, E.M.;

Eur. J. Immunol. 22, 491-498, 1992

A:Title: Mouse autoreactive gamma/delta T cells. II. Molecular characterization of the T

A:Reference number: A49037; MUID:92164730; PMID:1311262

A:Accession: C49037

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-11 <EQZ>

A:Cross-references: GB:S90639; NID:G246292; PIDN:AAB21549.1; PID:G246293

A:Experimental source: dendritic epidermal T-cell lines

A>Note: sequence extracted from NCBI backbone (NCBIN:90639, NCBIP:90645)

Query Match 26.5%; Score 22; DB 2; Length 11;

Best Local Similarity 66.7%; Pred. No. 1.3e+03;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 11 VVMGSGT 16

||: |||

DB 3 VVMGGT 8

RESULT 3

A47421
leukotriene B-4 12-hydroxydehydrogenase (EC 1.1.1.1.-) - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 19-May-1995 #sequence_revision 19-May-1995 #text_change 17-Mar-1999
C:Accession: A47421
R.Yokomizo, T.; Izumi, T.; Takahashi, T.; Kasama, T.; Kobayashi, Y.; Sato, F.; Taketani, J. Biol. Chem. 268, 18128-18135, 1993
A>Title: Enzymatic inactivation of leukotriene B-4 by a novel enzyme found in the porcine leukotriene synthase gene
A:Reference number: A47421; PMID:93352633; PMID:8394361
A:Accession: A47421
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-14 <YOK>
C:Keywords: oxidoreductase

Query Match 26.5%; Score 22; DB 2; Length 14;
Best Local Similarity 42.9%; Pred. No. 1.6e+03;
Matches 6; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 2 VIAKQMTYKVYMSG 15
DB 1 VRKSWTLKKHPVG 14

RESULT 4
PH1616
IG H chain V-D-J region (clone B-less 30) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 17-Mar-1999
C:Accession: PH1616
R.Levinson, D.A.; Campos-Torres, J.; Leder, P.
J. Exp. Med. 178, 317-329, 1993
A>Title: Molecular characterization of transgene-induced immunodeficiency in B-less mice
A:Reference number: PH1580; PMID:93301609; PMID:8315387
A:Accession: PH1616
A:Molecule type: DNA
A:Residues: 1-15 <LEV>
A:Experimental source: bone marrow pre-B lymphocyte
C:Keywords: immunoglobulin

Query Match 26.5%; Score 22; DB 2; Length 15;
Best Local Similarity 57.1%; Pred. No. 1.7e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 8 TYKVYMS 14
DB 7 TYSNYLS 13

RESULT 5
B45895
T-cell surface glycoprotein CD28 short form - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Feb-1994
C:Accession: B45895
R.Lee, K.P.; Taylor, C.; Petryniak, B.; Turka, L.A.; June, C.H.; Thompson, C.B.
J. Immunol. 145, 344-352, 1990
A>Title: The genomic organization of the CD28 gene. Implications for the regulation of T cell activation
A:Reference number: B45895; PMID:90293482; PMID:2162892
A:Accession: B45895
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-16 <LEE>
A:Cross-references: GB:M37813
C:Keywords: glycoprotein

Query Match 26.5%; Score 22; DB 2; Length 16;
Best Local Similarity 28.6%; Pred. No. 1.8e+03;
Matches 4; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
QY 4 AKQMTYKVYMSGTV 17
DB 1 AVNLSYNEKNGTI 14

RESULT 6
S78756
ribosomal protein MRP-L5, mitochondrial - bovine (fragments)
C:Species: Bos primigenius taurus (cattle)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C:Accession: S78756
R.Graack, H.R.
submitted to the Protein Sequence Database, May 1999
A:Reference number: S78756
A:Accession: S78756
A:Molecule type: protein
A:Residues: 1-12;13-17 <GRA>
C:Keywords: mitochondrion

Query Match 26.5%; Score 22; DB 2; Length 17;
Best Local Similarity 55.6%; Pred. No. 1.9e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 VIAKQMTYK 10
DB 9 VEAKXLIYK 17

RESULT 7
C49215
urease (EC 3.5.1.5) large subunit UreB - Helicobacter felis (fragment)
C:Species: Helicobacter felis
C>Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 06-Jan-2003
C:Accession: C49215
R.Turbett, G.R.; Hoj, P.B.; Horne, R.; Mee, B.J.
Infect. Immun. 60, 5259-5266, 1992
A>Title: Purification and characterization of the urease enzymes of Helicobacter species
A:Reference number: A49215; PMID:93084378; PMID:1452359
A:Accession: C49215
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-12 <TUR>
A:Experimental source: ATCC 49179
A>Note: sequence extracted from NCBI backbone (NCBI:119484)
C:Superfamily: urease, alpha subunit; urease 62K chain homology
C:Keywords: hydrolase

Query Match 25.3%; Score 21; DB 2; Length 12;
Best Local Similarity 40.0%; Pred. No. 2.1e+03;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 5 KQMTYKVYMS 14
DB 2 KKISKEVYS 11

RESULT 8
S09732
photosystem I protein psal - spinach chloroplast (fragment)
C:Species: chloroplast Spinacia oleracea (spinach)
C>Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 16-Feb-1997
C:Accession: S09732
R.Ikeuchi, M.; Hirano, A.; Hiyama, T.; Inoue, Y.
FEBS Lett. 263, 274-278, 1990
A>Title: Polypeptide composition of higher plant photosystem I complex. Identification
A:Reference number: S09730; PMID:90242987; PMID:2185953
A:Accession: S09732
A:Molecule type: protein
A:Residues: 1-16 <IKE>
C:Genetics:
A:Gene: psal
A:Genome: chloroplast
C:Keywords: chloroplast; photosynthesis; photosystem I; transmembrane protein

Query Match 25.3%; Score 21; DB 2; Length 16;
Best Local Similarity 50.0%; Pred. No. 2.8e+03;

Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 9 YKVMYS 14
: || : ||
Db 4 FKTYLS 9

RESULT 9
B42965
talin (glycosylated sites) - chicken (fragment)
C:Species: Gallus gallus (chicken)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 28-Apr-1995
C:Accession: B42965
R:Hagmann, J.; Grob, M.; Burger, M.M.
J. Biol. Chem. 267, 14424-14428, 1992
A:Title: The cytoskeletal protein talin is O-glycosylated.
A:Reference number: A42965; MUID:92332560; PMID:1629228
A:Accession: B42965
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-17 <HAG>
A:Experimental source: gizzard
A>Note: sequence extracted from NCBI backbone (NCBIP:108592)

Query Match 25.3%; Score 21; DB 2; Length 17;
Best Local Similarity 42.9%; Pred. No. 2.9e+03;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 VIAKQMT 8
: || : ||
Db 2 ILANQLT 8

RESULT 10
S69164
ferredoxin a1 - Japanese radish (fragments)
C:Species: Kaiware daikon (Japanese radish)
C:Date: 10-Mar-1998 #sequence_revision 10-Mar-1998 #text_change 17-Apr-1998
C:Accession: S69164
R:Obata, S.; Nishimura, M.; Nagai, K.; Sakihama, N.; Shin, M.
Arch. Biochem. Biophys. 316, 797-802, 1995
A:Title: Four ferredoxins from Japanese radish leaves.
A:Reference number: S69164; MUID:95168867; PMID:7864635
A:Accession: S69164
A:Molecule type: protein
A:Residues: 1-17 <OBA>
C:Keywords: 2Fe-2S; electron transfer; iron-sulfur protein

Query Match 25.3%; Score 21; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 2.9e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 TYKVV 11
: || : ||
Db 2 TVKV 5

RESULT 11
S63493
dissimilatory sulfite reductase gamma chain, membrane-bound and soluble - Desulfovibrio
C:Species: Desulfovibrio desulfuricans
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C:Accession: S63493; S63494
R:Steuber, J.; Arendsen, A.F.; Hagen, W.R.; Kroneck, P.M.H.
Eur. J. Biochem. 233, 873-879, 1995
A:Title: Molecular properties of the dissimilatory sulfite reductase from Desulfovibrio
A:Reference number: S63489; MUID:96085152; PMID:8521853
A:Accession: S63493
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <STE>
A:Accession: S63494
A>Status: preliminary

A:Molecule type: protein
A:Residues: 1-8 <ST2>

Query Match 24.1%; Score 20; DB 2; Length 8;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 QMTYK 10
: || : ||
Db 2 EITYK 6

RESULT 12
G41946
T-cell receptor gamma chain (2t.23) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C:Accession: G41946
R:Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma gene
A:Reference number: A41946; MUID:92049316; PMID:1658619
A:Accession: G41946
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-9 <WHE>
C:Keywords: T-cell receptor

Query Match 24.1%; Score 20; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 TYKVYMSG 15
: || : ||
Db 1 SYGSYSSG 8

RESULT 13
G85802
hypothetical protein Z2947 [imported] - Escherichia coli (strain O157:H7, substrain EDL93;
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: G85802
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: G85802
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-9 <STO>
A:Cross-references: GB:AE005174; NID:g12515957; PIDN:AA056983.1; GSPDB:GN00145; UWGP:Z294;
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: Z2947

Query Match 24.1%; Score 20; DB 2; Length 9;
Best Local Similarity 50.0%; Pred. No. 2.8e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 MTYKVYMS 14
: || : ||
Db 1 MTYTFMLS 8

RESULT 14
S66458
ferredoxin - Rhizobium meliloti (fragment)
C:Species: Rhizobium meliloti
C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C:Accession: S66458
R:Riedel, K.U.; Jouanneau, Y.; Masepohl, B.; Puehler, A.; Klipp, W.

Eur. J. Biochem. 231, 742-746, 1995
 A:Title: A Rhizobium meliloti ferredoxin (FdXN) purified from Escherichia coli donates e
 A:Reference number: S66458; MUID:95377307; PMID:7649175
 A:Accession: S66458
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-10 <RIE>
 C:Genetics:
 A:Gene: fdxN

Query Match 24.1%; Score 20; DB 2; Length 10;
 Best Local Similarity 57.1%; Pred. No. 2.6e+03;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 VIAQMT 8
 :|||
 Db 4 IIASQXT 10

RESULT 15
 H41946
 T-cell receptor gamma chain (5t.12) - mouse (fragment)
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
 C:Accession: H41946
 R:Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
 Mol. Cell. Biol. 11, 5902-5909, 1991
 A:Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
 A:Reference number: A41946; MUID:92049316; PMID:1658619
 A:Accession: H41946
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-12 <WHE>
 C:Keywords: T-cell receptor

Query Match 24.1%; Score 20; DB 2; Length 12;
 Best Local Similarity 80.0%; Pred. No. 3.1e+03;
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 11 VYMSG 15
 :|||
 Db 7 VYSSG 11

Search completed: August 12, 2004, 07:05:39
 Job time : 38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 06:55:55 ; Search time 32 Seconds
(without alignments)
27.662 Million cell updates/sec

Title: US-09-890-463-2

Sequence: 83
1 SVIAKQMTYKYVMSGTV 17

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 943

Minimum DB seq length: 0
Maximum DB seq length: 17

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	22	26.5	9	1	P31929 homo sapien
2	20	24.1	15	1	MP2A ORYSA
3	20	24.1	15	1	PLAS_MICAE
4	19	22.9	14	1	ADFA_TENMO
5	19	22.9	15	1	MALT_BACTQ
6	19	22.9	16	1	PAZI_TRIST
7	18	21.7	13	1	ADFB_TENMO
8	18	21.7	15	1	ACEA_ACICA
9	18	21.7	16	1	ODPB_SOLTU
10	17	20.5	11	1	ES1_RAT
11	17	20.5	15	1	COXI_THUOB
12	16	19.3	7	1	WNA2_ACHFU
13	16	19.3	8	1	CPD1_ENTFA
14	16	19.3	11	1	NXSN_PSETE
15	16	19.3	11	1	UXB2_YEAST
16	16	19.3	13	1	CRBL_VESLA
17	16	19.3	13	1	CRBL_VESXA
18	16	19.3	14	1	SMS1_MYOSC
19	16	19.3	14	1	SMS_ALIMI
20	16	19.3	15	1	GTS_ASADI
21	16	19.3	15	1	NXSQ_PSETE
22	16	19.3	15	1	PLAC_SHEEP
23	16	19.3	15	1	UC27_MAIZE
24	16	19.3	16	1	HBD_CLOPA
25	16	19.3	17	1	TL09_SPIOL
26	15	18.1	8	1	AKH_TABAT
27	15	18.1	8	1	HTF2_PERAM
28	15	18.1	8	1	RT34_BOVIN
29	15	18.1	9	1	NEUX_HUMAN
30	15	18.1	10	1	HTF2_CARMO
31	15	18.1	10	1	HTF2_HELZE
32	15	18.1	10	1	HTF2_TABAT
33	15	18.1	12	1	CALM_TETH

34 15 18.1 12 1 PPK4_PERFU p82690 periplaneta
35 15 18.1 14 1 MAST_VESMA p04205 vespa manda
36 15 18.1 14 1 PH1_PRUSE p29263 prunus sero
37 15 18.1 15 1 GR78_HORSE p16392 equus cabal
38 15 18.1 15 1 LEC3_AXIPO p28588 axinella ol
39 15 18.1 15 1 R13A_SPIOL p2454 spinacia ol
40 15 18.1 15 1 RGGG_CARCR p21586 caretta car
41 15 18.1 15 1 UC30_MAIZE p80636 zea mays m
42 15 18.1 16 1 AU26_LIFRA p82393 litoria ran
43 15 18.1 16 1 BRB_BASAL p83187 basella alb
44 15 18.1 16 1 NMPX_SOLTU p80501 solanum tub
45 15 18.1 16 1 PH2_PRUSE p29264 prunus sero

ALIGNMENTS

RESULT 1
ULAD_HUMAN STANDARD; PRT; 9 AA.
AC P31929, 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Unknown protein from 2D-page of liver tissue (Spot 106) (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.

RC TISSUE=Liver;
RX MEDLINE=94147969; PubMed=8313870;
RA Hughes G.J., Fruiger S., Faquet N., Pasquali C., Sanchez J.-C.,
RA Tissot J.-D., Bairoch A., Appel R.D., Hochstrasser D.F.;
RT "Human liver protein map: update 1993.";
RL Electrophoresis 14:1216-1222(1993).
CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
protein is: 6, its MW is: 15 kDa.
DR SWISS-2DPAGE; P31929; HUMAN.
FT NON TER 9
SQ SEQUENCE 9 AA; 1129 MW; D02DFB41B6D33322 CRC64;
Query Match 26.5%; Score 22; DB 1; Length 9;
Best Local Similarity 33.3%; Pred No. 1.4e+05;
Matches 3; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 IAKQMTYKYV 11
: : : : :
Db 1 LVKQTYHI 9

RESULT 2
MP2A ORYSA STANDARD; PRT; 15 AA.
AC P83466;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Pollen allergen Ory s 2-A (Fragment).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE.
RC STRAIN=cv. Japonica; TISSUE=Pollen;
RA Kerim T., Imin N., Weinman J.J., Rolfe B.G.;
RL Submitted (SEP-2002) to Swiss-Prot.
CC -!- ALLERGEN: Causes an allergic reaction in human. Causes grass
pollen allergy. Binds IGE.
CC -!- SIMILARITY: Belongs to the expansin family.

```

DR InterPro; IPR007112; Expan_endogl.
DR PROSITE; PS00842; EXPANSIN_EG45; FALSE_NEG.
KW Allergen.
FT NON TER 15 15
SQ SEQUENCE 15 AA; 1513 MW; 2C65C9FB3632A1C CRC64;

Query Match 24.1%; Score 20; DB 1; Length 15;
Best Local Similarity 50.0%; Pred. No. 2.1e+03;
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 6 QMTYKV 11
Db 2 EVTFKV 7

RESULT 3
PLAS MICAE STANDARD; PRT; 15 AA.
ID _PLAS MICAE STANDARD; PRT; 15 AA.
AC P10625;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Plastocyanin (Fragment).
GN PETE.
OS Microcystis aeruginosa.
OC Bacteria; Cyanobacteria; Chroococcales; Microcystis.
OX NCBI_TaxID=1126;
[1]
RN RP SEQUENCE.
RX MEDLINE=89134784; PubMed=2537099;
RA Tan S., Ho K.-K.;
RT "Purification of an acidic plastocyanin from Microcystis aeruginosa.";
RL Biochim. Biophys. Acta 973:111-117(1989).
CC -!- FUNCTION: Participates in electron transfer between P700 and the
CC cytochrome b6-f complex in photosystem I.
CC -!- SIMILARITY: Contains 1 plastocyanin-like domain.
DR PIR; S03353; S03353.
DR HAWAP; MF_00566; -; 1.
DR InterPro; IPR000923; BlueCu 1.
DR PROSITE; PS00196; COPPER BLUE; PARTIAL.
KW Electron transport; Copper.
FT NON TER 15 15
SQ SEQUENCE 15 AA; 1555 MW; 32BD4662F4F969 CRC64;

Query Match 24.1%; Score 20; DB 1; Length 15;
Best Local Similarity 50.0%; Pred. No. 2.1e+03;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 8 TYKYVMSG 15
Db 2 TFTVRMGG 9

RESULT 4
ADFA_TENMO STANDARD; PRT; 14 AA.
ID _ADFA_TENMO STANDARD; PRT; 14 AA.
AC P82965;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Antidiuretic factor A (ADFA) (ADF) (Antidiuretic hormone A) (ADHA).
OS Tenebrio molitor (Yellow mealworm).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
OC Tenebrionidae; Tenebrio.
OX NCBI_TaxID=7067;
[1]
RN RP SEQUENCE, FUNCTION, TISSUE SPECIFICITY, AND MASS SPECTROMETRY.
RC TISSUE=Head;
RX MEDLINE=21642653; PubMed=11756661;
RA Eigenheer R.A., Nicolson S.W., Schegg K.M., Hull J.J., Schooley D.A.;
RT "Identification of a potent antidiuretic factor acting on beetle
Malpighian tubules.";

```

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:84-89(2002).
CC -!- FUNCTION: Strong inhibitor of fluid secretion by the Malpighian
CC tubules. Uses cGMP as a second messenger and inhibits fluid
CC production by decreasing cAMP concentration.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- MASS SPECTROMETRY: MW=1541.58; METHOD=MALDI.
CC -!- SIMILARITY: STRONG, TO THE C-TERMINAL OF T.MOLLITOR CUTICULAR
CC PROTEIN LPCP29.
DR GO; GO:0005576; C:extracellular; NAS.
DR GO; GO:0005184; F:neuropeptide hormone activity; NAS.
DR GO; GO:0007218; F:neuropeptide signaling pathway; NAS.
KW Neuropeptide; Hormone.
SQ SEQUENCE 14 AA; 1543 MW; F49C91A3F16E43D1 CRC64;

Query Match 22.9%; Score 19; DB 1; Length 14;
Best Local Similarity 50.0%; Pred. No. 3e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 7 MTKYKV 12
Db 9 VSYHYV 14

RESULT 5
MALT BACTQ STANDARD; PRT; 15 AA.
ID _MALT BACTQ STANDARD; PRT; 15 AA.
AC P80072;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Maltase (EC 3.2.1.20) (Alpha-glucosidase I) (Fragment).
OS Bacillus thermoamyloliquefaciens.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1425;
[1]
RN RP SEQUENCE.
RC STRAIN=KPI071 / FERM P8477;
RX MEDLINE=92209510; PubMed=1555585;
RA Suzuki Y., Yonezawa K., Hattori M., Takii Y.;
RT "Assignment of Bacillus thermoamyloliquefaciens KPI071
RT alpha-glucosidase I to an exo-alpha-1,4-glucosidase, and its striking
RT similarity to bacillary oligo-1,6-glucosidases in N-terminal sequence
RT and in structural parameters calculated from the amino acid
RT composition.";
RL Eur. J. Biochem. 205:249-256(1992).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing 1,4-
CC linked D-glucose residues with release of D-glucose.
CC -!- SIMILARITY: Belongs to family 13 of glycosyl hydrolases.
DR PIR; S21240; S21240.
KW Hydrolase; Glycosidase.
FT NON TER 15 15
SQ SEQUENCE 15 AA; 1929 MW; 62B4CE501F2D3042 CRC64;

Query Match 22.9%; Score 19; DB 1; Length 15;
Best Local Similarity 33.3%; Pred. No. 3.2e+03;
Matches 2; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 7 MTKYKV 12
Db 10 VVYQIV 15

RESULT 6
PA21_TRIST STANDARD; PRT; 16 AA.
ID _PA21_TRIST STANDARD; PRT; 16 AA.
AC P82892;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phospholipase A2, basic 1 (EC 3.1.1.4) (PA2-I) (PA2-I)
DE (Phosphatidylcholine 2-acylhydrolase) (Fragment).
OS Trimeresurus stejnegeri (Chinese green tree viper).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroides;
 OC Viperidae; Crotalinae; Trimersurus.
 OX NCBI_TaxID=39682;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Venom;
 RA Li S.-Y., Guo Z.-X., Yang Y.-Y., Wang W.-Y., Xiong Y.-L.;
 RT "Isolation and sequencing of five variants of phospholipase A2
 from venom of snake *Trimersurus stejnegeri*.";
 RL J. Hubei Univ. 25:63-68(2003).
 CC -!- FUNCTION: PA2 catalyzes the calcium-dependent hydrolysis of the 2-
 acyl groups in 3-sn-phosphoglycerides.
 CC -!- CATALYTIC ACTIVITY: Phosphatidylcholine + H(2)O = 1-
 acylglycerophosphocholine + a fatty acid anion.
 CC -!- COFACTOR: Binds 1 calcium ion per subunit (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -!- SIMILARITY: Belongs to the phospholipase A2 family. Group II
 subfamily.
 DR InterPro: IPR001211; PhospholipaseA2.
 DR PROSITE; PS00119; PA2_ASP; PARTIAL.
 DR PROSITE; PS00118; PA2_HIS; PARTIAL.
 KW Hydrolase; Lipid degradation; Calcium; Multigene family.
 FT NON TER 16
 SQ SEQUENCE 16 AA; 2012 MW; 4EF2D4959E981117 CRC64;
 Query Match 22.9%; Score 19; DB 1; Length 16;
 Best Local Similarity 66.7%; Pred. No. 3.4e+03;
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 5 KQMTYK 10
 DB 10 KQMTNK 15
 RESULT 7
 ADPB TENNO
 ID _ADPB TENNO STANDARD; PRT; 13 AA.
 AC P83109;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Antidiuretic factor B (ADFB).
 OS Tenebrio molitor (Yellow mealworm).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 OC Tenebrionidae; Tenebrio.
 OX NCBI_TaxID=7067;
 RN [1]
 RP SEQUENCE, FUNCTION, TISSUE SPECIFICITY, MASS SPECTROMETRY, AND
 RP SYNTHESIS.
 RC TISSUE=Head;
 RX MEDLINE=22465067; PubMed=12576082;
 RA Eigenheer R.A., Wiehart U.M., Nicolson S.W., Schoofs L., Schegg K.M.,
 Hull J.J., Schooley D.A.;
 RT "Isolation, identification and localization of a second beetle
 antidiuretic peptide.";
 RL Peptides 24:27-34(2003).
 CC -!- FUNCTION: Inhibitor of fluid secretion by Malpighian tubules. Uses
 cGMP as second messenger. May function as an antidiuretic
 hormone.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Brain. Appears to be mainly expressed in two
 pairs of bilaterally symmetrical cells in the protocerebrum.
 CC -!- MASS SPECTROMETRY: MW=1560.33; METHOD=MALDI.
 KW Neuropeptide; Hormone.
 SQ SEQUENCE 13 AA; 1562 MW; 0240A4504B8A632B CRC64;
 Query Match 21.7%; Score 18; DB 1; Length 13;
 Best Local Similarity 37.5%; Pred. No. 4.2e+03;
 Matches 3; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 QY 8 TVKVMWSG 15

Db 5 SYKPHIYG 12
 :|| : :|
 RESULT 8
 ACEA ACICA STANDARD; PRT; 15 AA.
 AC P28467;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Isocitrate lyase (EC 4.1.3.1) (Isocitrase) (Isocitratase) (ICL)
 (Fragment).
 DE (Fragment).
 GN ACEA
 OS Acinetobacter calcoaceticus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Moraxellaceae; Acinetobacter.
 OX NCBI_TaxID=471;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=92041568; PubMed=1938889;
 RA Hoyt J.C., Johnson K.E., Reeves H.C.;
 RT "Purification and characterization of Acinetobacter calcoaceticus
 isocitrate lyase.";
 RL J. Bacteriol. 173:6844-6848(1991).
 CC -!- CATALYTIC ACTIVITY: Isocitrate = succinate + glyoxylate.
 CC -!- PATHWAY: Glyoxylate bypass; first step.
 CC -!- SUBUNIT: Homotetramer (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to the isocitrate lyase family.
 DR PIR; A41338; A41338.
 DR InterPro: IPR000918; Isocit lyase ph.
 DR PROSITE; PS00161; ISOCITRATE_LYASE; PARTIAL.
 KW Glyoxylate bypass; Tricarboxylic acid cycle; Lyase.
 FT NON TER 15
 SQ SEQUENCE 15 AA; 1710 MW; 83AE726B1F2F96E3 CRC64;
 Query Match 21.7%; Score 18; DB 1; Length 15;
 Best Local Similarity 75.0%; Pred. No. 4.8e+03;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 7 MTYK 10
 DB 1 MTYQ 4
 :|| : :|
 RESULT 9
 ODPB SOLTU
 ID _ODPB SOLTU STANDARD; PRT; 16 AA.
 AC P81419;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pyruvate dehydrogenase E1 component beta subunit, mitochondrial
 (EC 1.2.4.1) (PDHE1-B) (Fragment).
 OS Solanum tuberosum (Potato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC Lamiales; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4113;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=cv. Romano; TISSUE=Tuber;
 RX MEDLINE=98399821; PubMed=9729464;
 RA Millar A.H., Knorr C., Leaver C.J., Hill S.A.;
 RT "Plant mitochondrial pyruvate dehydrogenase complex: purification and
 identification of catalytic components in potato.";
 RL Biochem. J. 334:571-576(1998).
 CC -!- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall
 conversion of pyruvate to acetyl-CoA and CO(2). It contains
 multiple copies of three enzymatic components: pyruvate
 dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and
 lipoamide dehydrogenase (E3).

```

CC -!- CATALYTIC ACTIVITY: Pyruvate + lipoamide = S-
CC acetylhydrolipoamide + CO(2).
CC -!- COPACTOR: Thiamine pyrophosphate.
CC -!- SUBUNIT: Tetramer of two alpha and two beta subunits (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC Glycolysis; Oxidoreductase; Flavoprotein; Thiamine pyrophosphate;
KW Phosphorylation; Mitochondrion.
FT NON_TER 16
SQ SEQUENCE 16 AA; 1705 MW; FF6ED80EC804F797 CRC64;

Query Match 21.7%; Score 18; DB 1; Length 16;
Best Local Similarity 50.0%; Pred. No. 5.1e+03;
Matches 3; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 5 KQMTYK 10
Db 5 KEMTVR 10

RESULT 10
ES1_RAT
ID ES1_RAT STANDARD; PRT; 11 AA.
AC P56571;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE R1 protein, mitochondrial (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RC STRAIN=Wistar; TISSUE=Heart;
RA Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,
RA Jungblut P.R.;
RL Submitted (SEP-1998) to Swiss-Prot.
CC -!- SUBCELLULAR LOCATION: Mitochondrial (Potential).
CC -!- MISCELLANEOUS: BY 2D-PAGE, the determined pI of this protein (spot
CC P2) is: 8.9, its MW is: 25 kDa.
CC -!- SIMILARITY: BELONGS TO THE ES1 FAMILY.
KW Mitochondrion.
FT NON_TER 11
SQ SEQUENCE 11 AA; 1142 MW; D862272D32C72DC2 CRC64;

Query Match 20.5%; Score 17; DB 1; Length 11;
Best Local Similarity 75.0%; Pred. No. 5.4e+03;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 13 MSGT 16
Db 8 LSGT 11

RESULT 11
COXI_THUOB
ID COXI_THUOB STANDARD; PRT; 15 AA.
AC P80978;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Cytochrome c oxidase polypeptide VIC-2 (EC 1.9.3.1) (Fragments).
OS Thunnus obesus (Bigeye tuna).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8241;
RN [1]
RP SEQUENCE.
RC TISSUE=Heart;
RA MEDLINE=97454291; PubMed=9310366;
RX

```

```

RA Arnold S., Lee I., Kim M., Song E., Linder D., Lottspeich F.,
RA Kadenbach B.;
RT "The subunit structure of cytochrome-c oxidase from tuna heart and
RT liver.";
RL Eur. J. Biochem. 248:99-103(1997).
CC -!- FUNCTION: This protein is one of the nuclear-coded polypeptide
CC chains of cytochrome c oxidase, the terminal oxidase in
CC mitochondrial electron transport.
CC -!- CATALYTIC ACTIVITY: 4 ferrocycochrome c + O(2) = 4 ferricytochrome
CC c + 2 H(2)O.
CC -!- SUBCELLULAR LOCATION: Mitochondrial inner membrane.
DR PIR; S77987; S77987.
KW Oxidoreductase; Inner membrane; Mitochondrion.
FT NON_TER 1
FT NON_CONS 8
FT NON_TER 15
SQ SEQUENCE 15 AA; 1696 MW; 4C4C966C73A40294 CRC64;

Query Match 20.5%; Score 17; DB 1; Length 15;
Best Local Similarity 60.0%; Pred. No. 7.2e+03;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 VIAKQ 6
Db 5 VVAKK 9

RESULT 12
WWA2_ACHFU
ID WWA2_ACHFU STANDARD; PRT; 7 AA.
AC P35920;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
DE Wwamide-2.
OS Achatina fulica (Giant African snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC Sigmurethra; Achatinoidea; Achatinidae; Achatina.
OX NCBI_TaxID=6530;
RN [1]
RP SEQUENCE.
RC TISSUE=Ganglion;
RX MEDLINE=93265912; PubMed=8495720;
RA Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT "Wwamide-1, -2 and -3: novel neuromodulatory peptides isolated from
RT ganglia of the African giant snail, Achatina fulica.";
RL FEBS Lett. 323:104-108(1993).
DR PIR; S33246; S33246.
KW Neuropeptide; Amidation.
FT MOD_RES 7
SQ SEQUENCE 7 AA; 964 MW; 7362D5B686D32310 CRC64;

Query Match 19.3%; Score 16; DB 1; Length 7;
Best Local Similarity 75.0%; Pred. No. 1.4e+05;
Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 KQMT 8
Db 2 KQMS 5

RESULT 13
CPD1_ENTFA
ID CPD1_ENTFA STANDARD; PRT; 8 AA.
AC P13269;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-FEB-1991 (Rel. 17, Last annotation update)
DE Sex pheromone CPD1.
OS Enterococcus faecalis (Streptococcus faecalis).
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=1351;
RN [1]

```

```

RP SEQUENCE.
RX MEDLINE=85040388; PubMed-6436978;
RA Suzuki A., Mori M., Sagakami Y., Isogai A., Fujino M., Kitada C.,
RA Craig R.A., Clewell D.B.;
RT "Isolation and structure of bacterial sex pheromone, cPD1.";
RL Science 226:849-850(1984).
CC -!- FUNCTION: cPD1 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC BACTERIOCIN PLASMID PPd1.
CC Pheromone.
KW SEQUENCE 8 AA; 913 MW; 8665B79C682C729 CRC64;
SQ SEQUENCE 8 AA; 913 MW; 8665B79C682C729 CRC64;

Query Match 19.3%; Score 16; DB 1; Length 8;
Best Local Similarity 40.0%; Pred. No. 1.4e+05;
Matches 2; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 11 VMSG 15
Db 4 MFLSG 8

RESULT 14
NXSN_PSETE STANDARD; PRT; 11 AA.
AC P59072;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Short neurotoxin.N1 (Alpha neurotoxin) (Fragment).
OS Pseudonaja textilis (Eastern brown snake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Elapidae; Acanthophiinae; Pseudonaja.
OX NCBI_TaxID=8673;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX MEDLINE=99449602; PubMed=10518793;
RA Gong N.L., Armugam A., Jeyaseelan K.;
RT "Postsynaptic short-chain neurotoxins from Pseudonaja textilis: CDNA
RT cloning, expression and protein characterization.";
RL Eur. J. Biochem. 265:982-989(1999).
CC -!- FUNCTION: Lethal neurotoxin, binds and inhibits nicotinic
CC acetylcholine receptors (nAChR).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
CC -!- MASS SPECTROMETRY: MW=6236; METHOD=Electrospray.
CC -!- MISCELLANEOUS: LD(50) is 0.84 mg/kg by intravenous injection.
CC -!- SIMILARITY: Belongs to the snake toxin family.
DR InterPro; IPR003571; Snake toxin.
DR PROSITE; PS00272; SNAKE_TOXIN; PARTIAL.
KW Toxin; Neurotoxin; Postsynaptic neurotoxin;
KW Acetylcholine receptor inhibitor; Multigene family.
FT UNSURE 3
FT NON_TER 11
SQ SEQUENCE 11 AA; 1319 MW; 0D1EF0C81B58732B CRC64;

Query Match 19.3%; Score 16; DB 1; Length 11;
Best Local Similarity 75.0%; Pred. No. 8.1e+03;
Matches 3; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 YKXY 12
Db 4 YKGY 7

RESULT 15
UXB2 YEAST STANDARD; PRT; 11 AA.
AC P99013;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Unknown protein from 2D-page (Spot 2D-000K2F) (Fragment).

```

```

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE.
RC STRAIN=X2180-1A;
RA Sanchez J.-C., Golaz O., Schaller D., Morch F., Frutiger S.,
RA Hughes G.J., Appel R.D., Deshusses J., Hochstrasser D.F.;
RL Submitted (AUG-1995) to Swiss-Prot.
CC -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC protein is: 6.20, its MW is: 9.2 kDa.
DR SWISS-2DPAGE; P99013; YEAST.
FT NON_TER 11
SQ SEQUENCE 11 AA; 1328 MW; EC38021C0DCB42DA CRC64;

Query Match 19.3%; Score 16; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 8.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 KVV 12
Db 9 KVV 11

```

Search completed: August 12, 2004, 07:04:05
 JOB time : 33 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 07:01:00 ; Search time 34 Seconds

(without alignments)
157.759 Million cell updates/sec

Title: US-09-890-463-2

Perfect score: 83

Sequence: 1 SVIAQMTYKVYMSGTV 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 5061

Minimum DB seq length: 0

Maximum DB seq length: 17

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriaph:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	23	27.7	13	Q8QGZ5	Q8qgz5 fugu rubrip
2	23	27.7	15	Q9XJL8	Q9xjl8 grus leucog
3	23	27.7	15	Q9TH03	Q9th03 grus paradi
4	22	26.5	13	Q8T6E9	Q8t6e9 drosophila
5	22	26.5	13	Q8STI5	Q8sti5 drosophila
6	22	26.5	13	Q35793	Q35793 saccharomyc
7	21	25.3	8	Q95213	Q95213 oryctolagus
8	21	25.3	12	Q9R3B3	Q9r3b3 helicobacte
9	21	25.3	13	Q8GL29	Q8gl29 borrelia bu
10	21	25.3	7	Q9R575	Q9r575 synecocyst
11	20	24.1	19	Q8X4G1	Q8x4g1 escherichia
12	20	24.1	11	Q77900	Q77900 oreochromis
13	20	24.1	11	Q77917	Q77917 oreochromis
14	20	24.1	11	Q77902	Q77902 oreochromis
15	20	24.1	11	Q77921	Q77921 pseudotroph
16	20	24.1	11	Q77901	Q77901 oreochromis

17	20	24.1	11	7	Q77916	Q77916 oreochromis
18	20	24.1	11	7	Q77905	Q77905 oreochromis
19	20	24.1	11	7	Q77899	Q77899 oreochromis
20	20	24.1	11	7	Q77904	Q77904 oreochromis
21	20	24.1	11	7	Q77903	Q77903 oreochromis
22	20	24.1	12	6	Q9BFT9	Q9bft9 tupaia mino
23	20	24.1	12	8	Q9GE05	Q9ge05 etaballia g
24	20	24.1	14	5	Q9NFK8	Q9nfk8 brugia paba
25	20	24.1	14	8	Q8M099	Q8m099 tockus nasu
26	20	24.1	14	10	Q7X8F6	Q7x8f6 zea mays (m
27	20	24.1	15	2	Q52059	Q52059 salmonella
28	20	24.1	15	2	Q05991	Q05991 staphylococ
29	20	24.1	15	10	Q7XB01	Q7xb01 zea mays (m
30	20	24.1	16	10	Q7XB00	Q7xb00 zea mays (m
31	20	24.1	17	6	Q95M49	Q95m49 bos taurus
32	20	24.1	17	10	Q7XAZ9	Q7xaz9 zea mays (m
33	19	22.9	8	2	P83158	P83158 anabaena sp
34	19	22.9	12	4	Q9BY99	Q9by99 homo sapien
35	19	22.9	12	6	Q9BFT6	Q9bft6 ateles fusc
36	19	22.9	12	6	Q9BEV6	Q9bev6 chaetophrac
37	19	22.9	12	6	Q9BFR6	Q9bfr6 canis famil
38	19	22.9	12	6	Q9BFS3	Q9bfs3 okapia john
39	19	22.9	12	6	Q9BFR1	Q9bfr1 condylura c
40	19	22.9	12	6	Q9BFT7	Q9bft7 tarsius ban
41	19	22.9	12	6	Q9BFR4	Q9bfr4 manis penta
42	19	22.9	12	6	Q9BFS9	Q9bfs9 megaptera n
43	19	22.9	12	6	Q9BFR3	Q9bfr3 erinaceus c
44	19	22.9	12	6	Q9BEV5	Q9bev5 tamandua te
45	19	22.9	12	6	Q9BFS5	Q9bfs5 tragelaphus

ALIGNMENTS

RESULT 1

Q8QGZ5 ID Q8QGZ5 PRELIMINARY; PRT; 13 AA.
AC Q8QGZ5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Guanine nucleotide binding protein (Fragment).
GN GNAO.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97129408; PubMed=8973916;
RA Sarwal M.M., Sontag J.M., Hoang L., Brenner S., Wilkie T.M.;
RT "G protein alpha subunit multigene family in the Japanese puffer fish
Fugu rubripes: PCR from a compact vertebrate genome.";
RL Genome Res. 6:1207-1215(1996).
DR ENBL; L79891; AAL7640.1;
DR GO; GO:0004871; F:signal transducer activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . ; IEA.
DR InterPro; IPR001019; Gprotein_alpha.
DR Pfam; PF00503; G-alpha; 1.
FT NON_TER 1 13
FT NON_TER 13 13
SQ SEQUENCE 13 AA; 1336 MW; 455B59640B44B5B3 CRC64;

Query Match 27.7%; Score 23; DB 13; Length 13;

Best Local Similarity 71.4%; Pred. No. 3.6e+03;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SVIAKOM 7

Db 7 STIVKOM 13

```

RESULT 2
Q9XLJ8
ID Q9XLJ8 PRELIMINARY; PRT; 15 AA.
AC Q9XLJ8;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE NADH dehydrogenase subunit 6 (Fragment).
GN ND6.
OS Grus leucogeranus (Siberian crane).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Gruiformes; Gruidae; Grus.
OX NCBI_TaxID=40819;
RN [1]
RP SEQUENCE FROM N.A.
RA Glenn T.C., Stephan W., Braun M.J.;
RT "Effects of a Population Bottleneck on Whooping Crane Mitochondrial
RT DNA Variation."
RL Conserv. Biol. 0:0-0(1999).
RL EMBL; AF112371; AAD23992.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 15
SQ SEQUENCE 15 AA; 1807 MW; A7FFB3A65E8A734F CRC64;

Query Match 27.7%; Score 23; DB 8; Length 15;
Best Local Similarity 44.4%; Pred. No. 4.2e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 MTYKVMMSG 15
DB 1 MTYFMFFMG 9

RESULT 3
Q9TH03
ID Q9TH03 PRELIMINARY; PRT; 15 AA.
AC Q9TH03;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE NADH dehydrogenase subunit 6 (Fragment).
GN ND6.
OS Grus paradisea (Blue crane).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Gruiformes; Gruidae; Grus.
OX NCBI_TaxID=40825;
RN [1]
RP SEQUENCE FROM N.A.
RA Glenn T.C., Stephan W., Braun M.J.;
RT "Effects of a Population Bottleneck on Whooping Crane Mitochondrial
RT DNA Variation."
RL Conserv. Biol. 0:0-0(1999).
RL EMBL; AF112372; AAD23993.1; -.
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 15
SQ SEQUENCE 15 AA; 1807 MW; A7FFB3A65E8A734F CRC64;

Query Match 27.7%; Score 23; DB 8; Length 15;
Best Local Similarity 44.4%; Pred. No. 4.2e+03;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 7 MTYKVMMSG 15
DB 1 MTYFMFFMG 9

RESULT 4
Q8T6E9

```

```

ID Q8T6E9 PRELIMINARY; PRT; 13 AA.
AC Q8T6E9;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Myocyte enhancing factor 2 (Fragment).
GN MEF2.
OS Drosophila miranda (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7229;
RN [1]
RP SEQUENCE FROM N.A.
RA Schaeffer S.W., Goetting-Minesky M.P., Kovacevic M., Peoples J.,
RA Graybill J.L., Miller J.M., Kim K., Nelson J.G., Anderson W.W.;
RT "Evolutionary genomics of inversions in Drosophila pseudoobscura:
RT Modes of selection."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF476817; AAL91815.1; -.
DR FlyBase; FBgn0062329; Dmir\Wef2.
FT NON_TER 13
SQ SEQUENCE 13 AA; 1313 MW; 15BF380B6BE05050 CRC64;

Query Match 26.5%; Score 22; DB 5; Length 13;
Best Local Similarity 36.4%; Pred. No. 5.4e+03;
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 7 MTYKVMMSGTV 17
DB 3 MSLIYPSGSM 13

RESULT 5
Q8STI5
ID Q8STI5 PRELIMINARY; PRT; 13 AA.
AC Q8STI5;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Myocyte enhancing factor 2 (Fragment).
GN MEF2.
OS Drosophila pseudoobscura (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7237;
RN [1]
RP SEQUENCE FROM N.A.
RA Schaeffer S.W., Goetting-Minesky M.P., Kovacevic M., Peoples J.,
RA Graybill J.L., Miller J.M., Kim K., Nelson J.G., Anderson W.W.;
RT "Evolutionary genomics of inversions in Drosophila pseudoobscura:
RT Modes of selection."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF476728; AAL91726.1; -.
DR EMBL; AF476729; AAL91727.1; -.
DR EMBL; AF476730; AAL91728.1; -.
DR EMBL; AF476731; AAL91729.1; -.
DR EMBL; AF476732; AAL91730.1; -.
DR EMBL; AF476733; AAL91731.1; -.
DR EMBL; AF476734; AAL91732.1; -.
DR EMBL; AF476735; AAL91733.1; -.
DR EMBL; AF476736; AAL91734.1; -.
DR EMBL; AF476737; AAL91735.1; -.
DR EMBL; AF476738; AAL91736.1; -.
DR EMBL; AF476739; AAL91737.1; -.
DR EMBL; AF476740; AAL91738.1; -.
DR EMBL; AF476741; AAL91739.1; -.
DR EMBL; AF476742; AAL91740.1; -.
DR EMBL; AF476743; AAL91741.1; -.
DR EMBL; AF476744; AAL91742.1; -.
DR EMBL; AF476745; AAL91743.1; -.

```

DR EMBL; AF476745; AAL91744.1; -
DR EMBL; AF476747; AAL91745.1; -
DR EMBL; AF476748; AAL91746.1; -
DR EMBL; AF476749; AAL91747.1; -
DR EMBL; AF476750; AAL91748.1; -
DR EMBL; AF476751; AAL91749.1; -
DR EMBL; AF476752; AAL91750.1; -
DR EMBL; AF476753; AAL91751.1; -
DR EMBL; AF476754; AAL91752.1; -
DR EMBL; AF476755; AAL91753.1; -
DR EMBL; AF476756; AAL91754.1; -
DR EMBL; AF476757; AAL91755.1; -
DR EMBL; AF476758; AAL91756.1; -
DR EMBL; AF476759; AAL91757.1; -
DR EMBL; AF476760; AAL91758.1; -
DR EMBL; AF476761; AAL91759.1; -
DR EMBL; AF476762; AAL91760.1; -
DR EMBL; AF476763; AAL91761.1; -
DR EMBL; AF476764; AAL91762.1; -
DR EMBL; AF476765; AAL91763.1; -
DR EMBL; AF476766; AAL91764.1; -
DR EMBL; AF476767; AAL91765.1; -
DR EMBL; AF476768; AAL91766.1; -
DR EMBL; AF476769; AAL91767.1; -
DR EMBL; AF476770; AAL91768.1; -
DR EMBL; AF476771; AAL91769.1; -
DR EMBL; AF476772; AAL91770.1; -
DR EMBL; AF476773; AAL91771.1; -
DR EMBL; AF476774; AAL91772.1; -
DR EMBL; AF476775; AAL91773.1; -
DR EMBL; AF476776; AAL91774.1; -
DR EMBL; AF476777; AAL91775.1; -
DR EMBL; AF476778; AAL91776.1; -
DR EMBL; AF476779; AAL91777.1; -
DR EMBL; AF476780; AAL91778.1; -
DR EMBL; AF476781; AAL91779.1; -
DR EMBL; AF476782; AAL91780.1; -
DR EMBL; AF476783; AAL91781.1; -
DR EMBL; AF476784; AAL91782.1; -
DR EMBL; AF476785; AAL91783.1; -
DR EMBL; AF476786; AAL91784.1; -
DR EMBL; AF476787; AAL91785.1; -
DR EMBL; AF476788; AAL91786.1; -
DR EMBL; AF476789; AAL91787.1; -
DR EMBL; AF476790; AAL91788.1; -
DR EMBL; AF476791; AAL91789.1; -
DR EMBL; AF476792; AAL91790.1; -
DR EMBL; AF476793; AAL91791.1; -
DR EMBL; AF476794; AAL91792.1; -
DR EMBL; AF476795; AAL91793.1; -
DR EMBL; AF476796; AAL91794.1; -
DR EMBL; AF476797; AAL91795.1; -
DR EMBL; AF476798; AAL91796.1; -
DR EMBL; AF476799; AAL91797.1; -
DR EMBL; AF476800; AAL91798.1; -
DR EMBL; AF476801; AAL91799.1; -
DR EMBL; AF476802; AAL91800.1; -
DR EMBL; AF476803; AAL91801.1; -
DR EMBL; AF476804; AAL91802.1; -
DR EMBL; AF476805; AAL91803.1; -
DR EMBL; AF476806; AAL91804.1; -
DR EMBL; AF476807; AAL91805.1; -
DR EMBL; AF476808; AAL91806.1; -
DR EMBL; AF476809; AAL91807.1; -
DR EMBL; AF476810; AAL91808.1; -
DR EMBL; AF476811; AAL91809.1; -
DR EMBL; AF476812; AAL91810.1; -
DR EMBL; AF476813; AAL91811.1; -
DR EMBL; AF476814; AAL91812.1; -
DR EMBL; AF476815; AAL91813.1; -
DR EMBL; AF476816; AAL91814.1; -
DR FlyBase; FBgn0062239; Dpse\Mef2.
NON_TER 1 1

FT NON_TER 13
SQ SEQUENCE 13 AA; 1313 MW; 15BF380B6B505050 CRC64;

Query Match 26.5%; Score 22; DB 5; Length 13;
Best Local Similarity 36.4%; Pred. NO. 5.4e+03;
Matches 4; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 7 MTYKVMSTGV 17
|:|:|:|:
Db 3 MSLIIPSGSM 13

RESULT 6
Q35793 PRELIMINARY; PRT; 13 AA.
ID Q35793
AC Q35793; (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Inside intron 4 (Fragment).
OS Saccharomyces cerevisiae (Baker's yeast).
OG Mitochondrion.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D273-10B;
RX MEDLINE=81069885; PubMed=6354986;
RA Bonitz S.G., Coruzzi G., Thalenfeld B., Tzagoloff A., Macino G.;
RT "Assembly of the mitochondrial membrane system: Structure and
RT nucleotide sequence of the gene coding for subunit 1 of yeast
RT cytochrome oxidase".
RL J. Biol. Chem. 255:11927-11941 (1980).
DR EMBL; V00694; CAA24065.1; -
DR GO; GO:0005739; C:mitochondrion; IEA.
KW Mitochondrion.
FT NON_TER 1 1
FT NON_TER 13 13
SQ SEQUENCE 13 AA; 1749 MW; 11437826A89945B CRC64;

Query Match 26.5%; Score 22; DB 8; Length 13;
Best Local Similarity 40.0%; Pred. NO. 5.4e+03;
Matches 4; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 4 AKQYKVM 13
|:|:|:|:
Db 1 SKLYMYNYM 10

RESULT 7
Q95213 PRELIMINARY; PRT; 8 AA.
ID Q95213
AC Q95213; (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-FEB-1997 (TrEMBLrel. 19, Last annotation update)
DE Germline DH (DF) gene (Fragment).
GN DF.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F-1/rmg;
RA Mage R.G., Chen H.-T., Alexander C.B., Chen F.F.;
RT "Rabbit DQ52 and DH Gene Rearrangements in Early B-cell Development".
RL Mol. Immunol. 0:0-0 (1996).
DR EMBL; U62585; AAB18735.1; -
FT NON_TER 1 1
FT NON_TER 8 8
SQ SEQUENCE 8 AA; 845 MW; 5CA861B5AB58677B CRC64;

Query Match 25.3%; Score 21; DB 7; Length 8;
 Best Local Similarity 50.0%; Pred. No. 1e+06;
 Matches 4; Conservative 1; Mismatches 0; Gaps 0;

QY 9 YKVMSTG 16

DB 1 YFGYSTGT 8

RESULT 8

Q9R3B3

ID Q9R3B3 PRELIMINARY; PRT; 12 AA.

AC Q9R3B3; 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)

DE Urease large subunit (Fragment).

OS Helicobacter felis, and

OS Helicobacter pylori (Campylobacter pylori).

OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

OC Helicobacteraceae; Helicobacter.

OX NCBI_TaxID=214, 210;

RN [1]

RP SEQUENCE.

RX MEDLINE=93084378; PubMed=1452359;

RA Turbett G.R., Hoj P.B., Horne R., Mee B.J.;

RT "Purification and characterization of the urease enzymes of

RT Helicobacter species from humans and animals.;"

RL Infect. Immun. 60:5259-5266(1992).

DR PIR: C49215; C49215.

SQ SEQUENCE 12 AA; 1500 MW; 93F05D2362132415 CRC64;

Query Match 25.3%; Score 21; DB 2; Length 12;

Best Local Similarity 40.0%; Pred. No. 7.5e+03;

Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 KQMTYKVM 14

DB 2 KKISKEYVS 11

RESULT 9

Q8GL29

ID Q8GL29 PRELIMINARY; PRT; 13 AA.

AC Q8GL29;

DT 01-MAR-2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)

DE Bdr protein (Fragment).

GN BDR.

OS Borrelia burgdorferi (Lyme disease spirochete).

OG Plasmid group cp32-1.

OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.

OX NCBI_TaxID=139;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Sh-2-82;

RA Stevenson B., Miller J.C.;

RT "Comparative analyses of Borrelia burgdorferi erp genes and their cp32

RT prophages: conservation amidst diversity.;"

RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY142089; AAN17872.1; -.

DR GO; GO:0046821; C:extrachromosomal DNA; IEA.

KW Plasmid.

FT NON TER

SQ SEQUENCE 13 AA; 1548 MW; 08E030E20078A32B CRC64;

Query Match 25.3%; Score 21; DB 2; Length 13;

Best Local Similarity 50.0%; Pred. No. 8.2e+03;

Matches 3; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 MTYKVY 12

Db 4 LAYKTY 9

RESULT 10

Q9R575

ID Q9R575 PRELIMINARY; PRT; 17 AA.

AC Q9R575;

DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)

DE NAD(P)H-plastoquinone-oxidoreductase 18 kDa polypeptide (Fragment).

OS Synchocystis.

OC Bacteria; Cyanobacteria; Chroococcales.

OX NCBI_TaxID=1142;

RN [1]

RP SEQUENCE.

RX MEDLINE=93314795; PubMed=8325373;

RA Berger S., Ellersiek U., Kinzelt D., Steinmuller K.;

RT "Immunopurification of a subcomplex of the NAD(P)H-plastoquinone-

RT oxidoreductase from the cyanobacterium Synchocystis sp. PCC6803.;"

RL FEBS Lett. 326:246-250(1993).

SQ SEQUENCE 17 AA; 1890 MW; F77D9E9FE2A58FF7 CRC64;

Query Match 25.3%; Score 21; DB 2; Length 17;

Best Local Similarity 40.0%; Pred. No. 1.1e+04;

Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 8 TYKVYMSGTV 17

DB 3 TVKVVLNETI 12

RESULT 11

Q8X4G1

ID Q8X4G1 PRELIMINARY; PRT; 9 AA.

AC Q8X4G1;

DT 01-MAR-2002 (Tremblrel. 20, Created)

DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)

DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)

DE Hypothetical protein Z2947.

GN Z2947.

OS Escherichia coli O157:H7.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=83334;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=O157:H7 / EDL933 / ATCC 700927;

RX MEDLINE=21074935; PubMed=11206551;

RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,

RA Rose D.J., Mayhew G.P., Evans P.S., Gregor J., Kirkpatrick H.A.,

RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,

RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,

RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,

RA Welch R.A., Blattner F.R.;

RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.;"

RL Nature 409:529-533(2001).

DR EMBL; AE005411; AAG56883.1; -.

DR PIR; G85802; G85802.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 9 AA; 1107 MW; 8F6CB72699D1BB41 CRC64;

Query Match 24.1%; Score 20; DB 16; Length 9;

Best Local Similarity 50.0%; Pred. No. 1e+06;

Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 7 MTYKVYMS 14

DB 1 MTYTFMLS 8

RESULT 12

```

O77900
ID O77900 PRELIMINARY; PRT; 11 AA.
AC O77900;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MHC class II B locus 14 (Fragment)
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98315113; PubMed=9649539;
RA Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
RA Figueroa F., Sultmann H., Klein J.;
RT "Linkage relationships and haplotype polymorphism among cichlid mhc
RT class II B loci.";
RL Genetics 149:1527-1537(1998).
DR EMBL; AF050010; AAC41349.1; -.
FT NON TER 1 1
FT NON TER 11 11
SQ SEQUENCE 11 AA; 1349 MW; 81C12D8EB7341B41 CRC64;

Query Match 24.1%; Score 20; DB 7; Length 11;
Best Local Similarity 60.0%; Pred. No. le+04;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 MTKV 11
DB 3 MTKV 7

RESULT 13
O77917
ID O77917 PRELIMINARY; PRT; 11 AA.
AC O77917;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MHC class II B locus 14 (Fragment)
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98315113; PubMed=9649539;
RA Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
RA Figueroa F., Sultmann H., Klein J.;
RT "Linkage relationships and haplotype polymorphism among cichlid mhc
RT class II B loci.";
RL Genetics 149:1527-1537(1998).
DR EMBL; AF050030; AAC41369.1; -.
FT NON TER 1 1
FT NON TER 11 11
SQ SEQUENCE 11 AA; 1349 MW; 81C12D8EB7341B41 CRC64;

Query Match 24.1%; Score 20; DB 7; Length 11;
Best Local Similarity 60.0%; Pred. No. le+04;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 MTKV 11
DB 3 MTKV 7

RESULT 14
O77902
ID O77902 PRELIMINARY; PRT; 11 AA.
AC O77902;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MHC class II B locus 14 (Fragment)
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98315113; PubMed=9649539;
RA Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
RA Figueroa F., Sultmann H., Klein J.;
RT "Linkage relationships and haplotype polymorphism among cichlid mhc
RT class II B loci.";
RL Genetics 149:1527-1537(1998).
DR EMBL; AF050012; AAC41351.1; -.
FT NON TER 1 1
FT NON TER 11 11
SQ SEQUENCE 11 AA; 1349 MW; 81C12D8EB7341B41 CRC64;

Query Match 24.1%; Score 20; DB 7; Length 11;
Best Local Similarity 60.0%; Pred. No. le+04;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 MTKV 11
DB 3 MTKV 7

RESULT 15
O77921
ID O77921 PRELIMINARY; PRT; 11 AA.
AC O77921;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE MHC class II B locus 14 (Fragment)
OS Pseudotropheus sp. 'Pseudotropheus tropheops complex'.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidae;
OC Cichlidae; Pseudotropheus.
OX NCBI_TaxID=51796;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98315113; PubMed=9649539;
RA Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
RA Figueroa F., Sultmann H., Klein J.;
RT "Linkage relationships and haplotype polymorphism among cichlid mhc
RT class II B loci.";
RL Genetics 149:1527-1537(1998).
DR EMBL; AF050034; AAC41373.1; -.
FT NON TER 1 1
FT NON TER 11 11
SQ SEQUENCE 11 AA; 1349 MW; 81C12D8EB7341B41 CRC64;

Query Match 24.1%; Score 20; DB 7; Length 11;
Best Local Similarity 60.0%; Pred. No. le+04;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 MTKV 11
DB 3 MTKV 7

Search completed: August 12, 2004, 07:04:51
Job time : 36 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 12, 2004, 06:55:24 ; Search time 50 Seconds

(Without alignments)
96.066 Million cell updates/sec

Title: US-09-890-463-2

Perfect score: 83

Sequence: 1 SVIAKQMTYKVMGTV 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 470470

Minimum DB seq length: 0
Maximum DB seq length: 17

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 29Jan04: *
1: Geneseqp1980s: *
2: Geneseqp1990s: *
3: Geneseqp2000s: *
4: Geneseqp2001s: *
5: Geneseqp2002s: *
6: Geneseqp2003as: *
7: Geneseqp2003bs: *
8: Geneseqp2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	83	100.0	17	3 AAY97148	Pigment p
2	73	88.0	16	5 ABB99066	N-termina
3	72	86.7	16	5 ABB99073	N-termina
4	70	84.3	16	5 ABB99072	N-termina
5	69	83.1	16	5 ABB99068	N-termina
6	69	83.1	16	5 ABB99070	N-termina
7	68	81.9	16	5 ABB99067	N-termina
8	67	80.7	16	5 ABB99071	N-termina
9	66	79.5	16	5 ABB99069	N-termina
10	62	74.7	16	5 ABB99074	N-termina
11	36	43.4	13	5 ABB70008	Colour Fa
12	29	34.9	14	2 AAR77526	p45 metal
13	29	34.9	14	2 AAW05846	Fusarium
14	29	34.9	15	5 AAM48968	Human zin
15	28	33.7	12	4 AAW39598	Human mel
16	28	33.7	12	4 AAB45642	Vasoactiv
17	28	33.7	13	2 AAR69362	Stearoyl-
18	28	33.7	13	4 AAB45641	Vasoactiv
19	28	33.7	13	4 AAB45639	Vasoactiv
20	28	33.7	13	5 AAB19614	Human ste
21	28	33.7	14	2 AAR79549	Analgesic
22	28	33.7	14	4 AAB88179	CD66 pept
23	28	33.7	14	4 AAB45638	Vasoactiv
24	28	33.7	14	4 AAB45622	Vasoactiv
25	28	33.7	15	2 AAR79548	Analgesic

26 28 33.7 15 4 AAB99955 Human lat
27 28 33.7 15 4 AAB45619 Vasoactiv
28 28 33.7 15 4 AAB45621 Vasoactiv
29 28 33.7 15 5 ABB71317 Human Sai
30 28 33.7 16 2 AAR79547 Analgesic
31 28 33.7 16 3 AAY85708 Peptide s
32 28 33.7 16 4 AAB45618 Vasoactiv
33 28 33.7 16 4 AAB45620 Vasoactiv
34 28 33.7 17 2 AAR79546 Analgesic
35 28 33.7 17 4 AAB45617 Vasoactiv
36 28 33.7 17 6 ADA90426 MS-Roche
37 28 33.7 17 6 ADA89996 Anti-Abet
38 27 32.5 8 4 ABB19278 HIV B62 s
39 27 32.5 8 4 ABB19181 HIV B62 s
40 27 32.5 9 4 ABB21094 HIV A03 m
41 27 32.5 9 4 ABB23330 HIV A11 m
42 27 32.5 9 4 ABB21373 HIV A03 m
43 27 32.5 9 4 ABB23124 HIV A11 m
44 27 32.5 9 7 ADD57378 HLA bindi
45 27 32.5 9 7 ADD57758 HLA bindi

ALIGNMENTS

RESULT 1

AAY97148
ID AAY97148 standard; peptide; 17 AA.

XX AC AAY97148;

XX DT 04-DEC-2000 (first entry)

XX DE Pigment protein from coral tissue N-terminal peptide 2.

XX KW N-terminal; pigment protein from coral tissue; PPCT; fluorescence;

XX KW tissue marker; fluorescent marker; dyestuff; sunscreen; ultra violet;

XX KW UV filter.

XX OS Acropora horrida.

XX PN WO200046233-A1.

XX PD 10-AUG-2000.

XX PF 02-FEB-2000; 2000WO-AU000056.

XX PR 02-FEB-1999; 99AU-00008463.

XX (UNSY) UNIV SYDNEY.

XX PI Hoegh-Guldberg O, Dove S;

XX DR WPI; 2000-532892/48.

XX PT Novel pigment protein derived from corals capable of emitting fluorescence upon irradiation by incident light useful as tissue marker, fluorescent marker or general dyestuff.

XX Claim 4; Page 42; 49pp; English.

XX CC The N-terminal peptides shown in AAY97147-48 are from pigment protein from coral tissue (PPCT). PPCT is capable of emitting fluorescence upon irradiation by incident light whose maximal absorbance is in the range of 300-700 nm. PPCT may be used as a tissue marker, fluorescent marker (e.g. to follow gene expression in transformed tissues) or general dyestuff (all claimed). PPCT may also be used in sunscreen formulations or UV filters (both claimed)

XX SQ Sequence 17 AA;

Query Match 100.0%; Score 83; DB 3; Length 17;

```
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SVIAKQMTYKVYMSGTV 17
Db 1 SVIAKQMTYKVYMSGTV 17

RESULT 2
ID ABB99066 standard; peptide; 16 AA.
XX ABB99066;
XX
DT 22-JAN-2003 (first entry)
XX
DE N-terminal amino acid sequence of a CPM #6.
XX
KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
XX UV sink; sunscreen.
XX
OS Unidentified.
XX
PN WO200270703-A2.
XX
PD 12-SEP-2002.
XX
PF 01-MAR-2002; 2002WO-GB000928.
XX
PR 02-MAR-2001; 2001US-0273227P.
PR 21-MAR-2001; 2001AU-00003874.
PR 15-OCT-2001; 2001US-0329816P.
XX
PA (NUFA-) NUFARM LTD.
PA (UYQU) UNIV QUEENSLAND.
PA (JONE/) JONES E L.
XX
PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
PI Hoegh-Guldberg IO, Prescott M;
XX
XX WPI; 2002-740765/80.
XX
PT Novel color-facilitating molecule for producing a biomatrix, has a
PT polypeptide which alone/along with molecules imparts altered visual
PT characteristics to cells in the absence of excitation by extraneous non-
PT white light.
XX
PS Claim 4; Page 280; 510pp; English.
XX
CC The invention relates to an isolated colour-facilitating molecule (CFM)
CC comprising a polypeptide which, in a cell, alone or together with one or
CC more other molecules imparts an altered visual characteristic to the cell
CC when visualised by a human eye in the absence of excitation by extraneous
CC non-white light or particle emission. CFMs are useful for producing a
CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
CC red coloured fleece. They are useful for producing coloured plant
CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
CC uses include transducing or intensifying an image, providing additional
CC light for growing phototropic organisms e.g. algae and/or corals, for
CC coating materials that experience UV damage e.g. plastics and car
CC upholstery. CFMs are useful in the flower industry, in the development of
CC new varieties of flowering plants. Other contemplated uses include,
CC expression markers, general reporter molecules, photon traps, UV sinks or
CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
CC fungal species, and in fruits and vegetables to enhance their
CC marketability. CFMs embedded in a gel matrix improve image quality in
CC situations of distorted light spectra (biomatrix). The first all-protein
CC chromophore to be isolated was Green Fluorescent protein (GFP). The
CC current sequence represents the N-terminal amino acid sequence of a
CC colour-facilitating molecule (CFM)
```

```
SQ Sequence 16 AA;
Query Match 88.0%; Score 73; DB 5; Length 16;
Best Local Similarity 93.8%; Pred. No. 2.2e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVIAKQMTYKVYMSGT 16
Db 1 SVIAKQMTYKVYMSGT 16

RESULT 3
ID ABB99073 standard; peptide; 16 AA.
XX ABB99073;
XX
DT 22-JAN-2003 (first entry)
XX
DE N-terminal amino acid sequence of a CFM #13.
XX
KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
XX UV sink; sunscreen.
XX
OS Unidentified.
XX
PN WO200270703-A2.
XX
PD 12-SEP-2002.
XX
PF 01-MAR-2002; 2002WO-GB000928.
XX
PR 02-MAR-2001; 2001US-0273227P.
PR 21-MAR-2001; 2001AU-00003874.
PR 15-OCT-2001; 2001US-0329816P.
XX
PA (NUFA-) NUFARM LTD.
PA (UYQU) UNIV QUEENSLAND.
PA (JONE/) JONES E L.
XX
PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;
PI Hoegh-Guldberg IO, Prescott M;
XX
XX WPI; 2002-740765/80.
XX
PT Novel color-facilitating molecule for producing a biomatrix, has a
PT polypeptide which alone/along with molecules imparts altered visual
PT characteristics to cells in the absence of excitation by extraneous non-
PT white light.
XX
PS Claim 4; Page 281; 510pp; English.
XX
CC The invention relates to an isolated colour-facilitating molecule (CFM)
CC comprising a polypeptide which, in a cell, alone or together with one or
CC more other molecules imparts an altered visual characteristic to the cell
CC when visualised by a human eye in the absence of excitation by extraneous
CC non-white light or particle emission. CFMs are useful for producing a
CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
CC red coloured fleece. They are useful for producing coloured plant
CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
CC uses include transducing or intensifying an image, providing additional
CC light for growing phototropic organisms e.g. algae and/or corals, for
CC coating materials that experience UV damage e.g. plastics and car
CC upholstery. CFMs are useful in the flower industry, in the development of
CC new varieties of flowering plants. Other contemplated uses include,
CC expression markers, general reporter molecules, photon traps, UV sinks or
CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
CC fungal species, and in fruits and vegetables to enhance their
CC marketability. CFMs embedded in a gel matrix improve image quality in
CC situations of distorted light spectra (biomatrix). The first all-protein
CC chromophore to be isolated was Green Fluorescent protein (GFP). The
```

CC current sequence represents the N-terminal amino acid sequence of a
CC colour-facilitating molecule (CFM)

XX SQ Sequence 16 AA;

Query Match 86.7%; Score 72; DB 5; Length 16;
Best Local Similarity 93.8%; Pred. No. 3.4e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SVIAKQMTYKVMST 16

Db 1 SVIAKQMTYKVMST 16

RESULT 4

ABB99072
ID ABB99072 standard; peptide; 16 AA.

XX AC ABB99072;

XX DT 22-JAN-2003 (first entry)

XX DE N-terminal amino acid sequence of a CFM #12.

XX KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
KW UV sink; sunscreen.

XX OS Unidentified.

XX PN WO200270703-A2.

XX PD 12-SEP-2002.

XX PF 01-MAR-2002; 2002WO-GB000928.

XX PR 02-MAR-2001; 2001US-0273227P.

XX PR 21-MAR-2001; 2001AU-00003874.

XX PR 15-OCT-2001; 2001US-0329816P.

XX PA (NUFA-) NUFARM LTD.

XX PA (UYOU) UNIV QUEENSLAND.

XX PA (JONE/) JONES E L.

XX PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;

XX PI Hoegh-Guldberg IO, Prescott M;

XX PI WPI; 2002-740765/80.

XX PS Claim 4; Page 281; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)
CC comprising a polypeptide which, in a cell, alone or together with one or
CC more other molecules imparts an altered visual characteristic to the cell
CC when visualised by a human eye in the absence of excitation by extraneous
CC non-white light or particle emission. CFMs are useful for producing a
CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
CC red coloured fleece. They are useful for producing coloured plant
CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
CC uses include transducing or intensifying an image, providing additional
CC light for growing phototropic organisms e.g. algae and/or corals, for
CC coating materials that experience UV damage e.g. plastics and car
CC upholstery. CFMs are useful in the flower industry, in the development of
CC new varieties of flowering plants. Other contemplated uses include,
CC expression markers, general reporter molecules, photon traps, UV sinks or
CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
CC fungal species, and in fruits and vegetables to enhance their

CC marketability. CFMs embedded in a gel matrix improve image quality in
CC situations of distorted light spectra (biomatrix). The first all-protein
CC chromophore to be isolated was Green Fluorescent protein (GFP). The
CC current sequence represents the N-terminal amino acid sequence of a
CC colour-facilitating molecule (CFM)

XX SQ Sequence 16 AA;

Query Match 84.3%; Score 70; DB 5; Length 16;
Best Local Similarity 93.8%; Pred. No. 7.9e-06;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SVIAKQMTYKVMST 16

Db 1 SVIAKQMTYKVMST 16

RESULT 5

ABB99068

ID ABB99068 standard; peptide; 16 AA.

XX AC ABB99068;

XX DT 22-JAN-2003 (first entry)

XX DE N-terminal amino acid sequence of a CFM #8.

XX KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
KW chromophore; biomatrix; transgenic animal; colouring agent;
KW flower industry; expression marker; reporter molecule; photon trap;
KW UV sink; sunscreen.

XX OS Unidentified.

XX PN WO200270703-A2.

XX PD 12-SEP-2002.

XX PF 01-MAR-2002; 2002WO-GB000928.

XX PR 02-MAR-2001; 2001US-0273227P.

XX PR 21-MAR-2001; 2001AU-00003874.

XX PR 15-OCT-2001; 2001US-0329816P.

XX PA (NUFA-) NUFARM LTD.

XX PA (UYOU) UNIV QUEENSLAND.

XX PA (JONE/) JONES E L.

XX PI Jones EL, Karan M, Brugliera F, Mason J, Dove SG;

XX PI Hoegh-Guldberg IO, Prescott M;

XX PI WPI; 2002-740765/80.

XX PS Claim 4; Page 280; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)
CC comprising a polypeptide which, in a cell, alone or together with one or
CC more other molecules imparts an altered visual characteristic to the cell
CC when visualised by a human eye in the absence of excitation by extraneous
CC non-white light or particle emission. CFMs are useful for producing a
CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
CC red coloured fleece. They are useful for producing coloured plant
CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
CC uses include transducing or intensifying an image, providing additional
CC light for growing phototropic organisms e.g. algae and/or corals, for
CC coating materials that experience UV damage e.g. plastics and car
CC upholstery. CFMs are useful in the flower industry, in the development of
CC new varieties of flowering plants. Other contemplated uses include,

CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC current sequence represents the N-terminal amino acid sequence of a
 CC colour-facilitating molecule (CFM)
 XX
 SQ Sequence 16 AA;

Query Match 83.1%; Score 69; DB 5; Length 16;
 Best Local Similarity 87.5%; Pred. No. 1.2e-05;
 Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SVIAKQMTYKYVMSGT 16
 |||||:|||||||
 Db 1 SVIATQVIKYVMSGT 16
 |||||:|||||||

RESULT 6
 ABB99070
 ID ABB99070 standard; peptide; 16 AA.
 XX
 AC ABB99070;

DT 22-JAN-2003 (first entry)
 XX
 DE N-terminal amino acid sequence of a CFM #10.

KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunsreen.

XX Unidentified.
 XX WO200270703-A2.

FN 12-SEP-2002.

PD 01-MAR-2002; 2002WO-GB000928.

XX 02-MAR-2001; 2001US-0273227P.

PR 21-MAR-2001; 2001AU-00003874.

PR 15-OCT-2001; 2001US-0329816P.

XX (NUFA-) NUFARM LTD.

PA (UYQU) UNIV QUEENSLAND.

PA (JONE/) JONES E L.

XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;

PI Hoegh-Guldberg IO, Prescott M;

XX WPI; 2002-740765/80.

XX Novel color-facilitating molecule for producing a biomatrix, has a
 PT polypeptide which alone/along with molecules imparts altered visual
 PT characteristics to cells in the absence of excitation by extraneous non-
 PT white light.

XX Claim 4; Page 281; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)
 CC comprising a polypeptide which, in a cell, alone or together with one or
 CC more other molecules imparts an altered visual characteristic to the cell
 CC when visualised by a human eye in the absence of excitation by extraneous
 CC non-white light or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant
 CC extracts, e.g. flavouring, beverage or juice or colouring agent. Other
 CC uses include transducing or intensifying an image, providing additional
 CC light for growing phototropic organisms e.g. algae and/or corals, for

CC coating materials that experience UV damage e.g. plastics and car
 CC upholstery. CFMs are useful in the flower industry, in the development of
 CC new varieties of flowering plants. Other contemplated uses include,
 CC expression markers, general reporter molecules, photon traps, UV sinks or
 CC in sunscreens. CFMs modify visible colour in edible and/or ornamental
 CC fungal species, and in fruits and vegetables to enhance their
 CC marketability. CFMs embedded in a gel matrix improve image quality in
 CC situations of distorted light spectra (biomatrix). The first all-protein
 CC chromophore to be isolated was Green Fluorescent protein (GFP). The
 CC current sequence represents the N-terminal amino acid sequence of a
 CC colour-facilitating molecule (CFM)
 XX
 SQ Sequence 16 AA;

Query Match 83.1%; Score 69; DB 5; Length 16;
 Best Local Similarity 87.5%; Pred. No. 1.2e-05;
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SVIAKQMTYKYVMSGT 16
 |||||:|||||||
 Db 1 SVIVTQMTYKYVMSGT 16
 |||||:|||||||

RESULT 7
 ABB99067
 ID ABB99067 standard; peptide; 16 AA.
 XX
 AC ABB99067;

DT 22-JAN-2003 (first entry)

XX N-terminal amino acid sequence of a CFM #7.

KW Colour facilitating molecule; CFM; green fluorescent protein; GFP;
 KW chromophore; biomatrix; transgenic animal; colouring agent;
 KW flower industry; expression marker; reporter molecule; photon trap;
 KW UV sink; sunsreen.

XX Unidentified.

XX WO200270703-A2.

XX 12-SEP-2002.

XX 01-MAR-2002; 2002WO-GB000928.

PR 02-MAR-2001; 2001US-0273227P.

PR 21-MAR-2001; 2001AU-00003874.

PR 15-OCT-2001; 2001US-0329816P.

XX (NUFA-) NUFARM LTD.

PA (UYQU) UNIV QUEENSLAND.

PA (JONE/) JONES E L.

XX Jones EL, Karan M, Brugliera F, Mason J, Dove SG;

PI Hoegh-Guldberg IO, Prescott M;

XX WPI; 2002-740765/80.

XX Novel color-facilitating molecule for producing a biomatrix, has a
 PT polypeptide which alone/along with molecules imparts altered visual
 PT characteristics to cells in the absence of excitation by extraneous non-
 PT white light.

XX Claim 4; Page 280; 510pp; English.

XX The invention relates to an isolated colour-facilitating molecule (CFM)
 CC comprising a polypeptide which, in a cell, alone or together with one or
 CC more other molecules imparts an altered visual characteristic to the cell
 CC when visualised by a human eye in the absence of excitation by extraneous
 CC non-white light or particle emission. CFMs are useful for producing a
 CC transgenic animal which exhibits a novel colour e.g. sheep with blue or
 CC red coloured fleece. They are useful for producing coloured plant